



Europäisches Patentamt
European Patent Office
Office européen des brevets

⑪ Publication number:

0 394 538
A1

⑫

EUROPEAN PATENT APPLICATION

⑬ Application number: 89107780.2

⑮ Int. Cl. 5: C12N 1/19, C12P 21/02,
C12N 15/81, // (C12N1/19,
C12R1:645)

⑭ Date of filing: 28.04.89

⑯ Date of publication of application:
31.10.90 Bulletin 90/44

⑰ Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE

⑲ Applicant: Rhein Biotech Gesellschaft für
biotechnologische Prozesse und Produkte
mbH
Erkrather Strasse 230
D-4000 Düsseldorf 1 (DE)

⑳ Inventor: Hollenberg, Cornelius P., Prof. Dr.
Chopin-Str. 7
D-4000 Düsseldorf (DE)
Inventor: Strasser, Alexander, Dr.
Ringelsweide 16
D-4000 Düsseldorf (DE)

㉑ Representative: Patentanwälte Grünecker,
Kinkeldey, Stockmair & Partner
Maximilianstrasse 58
D-8000 München 22 (DE)

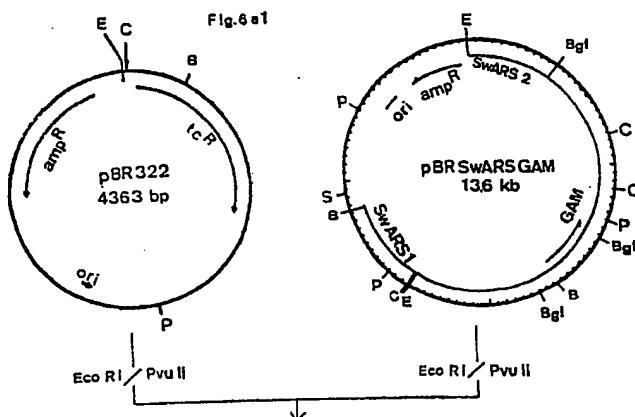
㉒ A yeast cell of the genus schwanniomyces.

㉓ The present invention relates to a yeast cell of the genus Schwanniomyces wherein said yeast cell contains at least one expression cassette comprising:

- a first DNA sequence serving as a regulon,
- optionally a second DNA sequence coding for a signal peptide,
- a third DNA sequence coding for a foreign protein, and
- optionally a fourth DNA sequence serving as a terminator.

In a preferred embodiment of the invention, the first, second and/or fourth DNA sequences are derived from genes coding for amylolytic enzymes of yeast cells of the genus Schwanniomyces. The use of these functional sequences in an expression cassette provides efficient expression of foreign genes in yeast cells according to the present invention and, furthermore, provides a system for efficient secretion of the expression product into the medium.

EP 0 394 538 A1



The present invention relates to yeast cells of the genus Schwanniomyces and a process for the production of polypeptides.

For expression of eukaryotic proteins of biotechnological relevance there exists a variety of host systems, which are selected in compliance with the specific requirements. The most frequently used host organism is Escherichia coli, being the best known representative of prokaryotic host organisms; Saccharomyces cerevisiae and animal tissue cultures are used as eukaryotic host cells. For certain purposes, e.g. where expression of a glycoprotein is desired, expression in E.coli or other prokaryotic organisms is not appropriate, since same lack any glycosylation capability. On the other hand, the mass production of protein in tissue cultures still is cumbersome and expensive; therefore, the organism of choice in many instances is yeast of the well characterized genus Saccharomyces. S. cerevisiae is able to secrete efficiently small proteins, e.g. polypeptides like MG1 (Kurjan and Herskowitz, 1982), β -endorphin or α -interferon (Bitter et al., 1984), however proteins of higher molecular weight are captured in the periplasmic space (Emr et al., 1981) or the cytoplasm, thus causing decreased growth rates or even death of the cell.

Therefore it is an object of the present invention to provide an improved eukaryotic expression system, which is optionally capable of secreting proteins encoded by a foreign DNA into the culture medium.

This object has been solved by a yeast cell of the genus Schwanniomyces, wherein said yeast cell contains at least one expression cassette comprising:

- a) a first DNA sequence serving as a regulon,
- b) optionally a second DNA sequence coding for a signal peptide
- c) a third DNA sequence coding for a foreign protein, and
- d) optionally a fourth DNA sequence serving as a terminator.

The invention, which comprises further subjects, is now described in a more detailed manner by the following description, examples and figures.

BRIEF DESCRIPTION OF THE FIGURES

Fig.1
A: Restriction map of the AMY1 gene encoding α -amylase and its 5' and 3' noncoding region
B: Nucleotide sequence of the AMY1 gene and its flanking region. In the 5' region the transcription initiation site is indicated by a dashed line above the sequence. Within the structural gene the presumptive signal peptidase cleavage sites are indicated by star-symbols. The two potential glycosylation sites, the cystein residues presumably forming four disulfide bonds (concluded by homology to the A. oryzae enzyme) are underlined. Within the 3' end region the transcription termination signal TAG...TATGT...TTT is indicated by a solid line.

Fig.2
A: Restriction map of the GAM1 gene encoding glucoamylase and its 5' and 3' noncoding region
B: Nucleotide sequence of the GAM1 gene and its flanking region. In the 5' region the transcription initiation site is indicated by a dashed line. Within the structural gene the presumptive signal peptides cleavage sites are indicated by star-symbols.

Fig.3
M13 Xhol. The EcoRI-Scal fragment, containing a 1.8 kb promoter fragment and approx. 0.4 kb structural gene fragment of α -amylase gene, was subcloned into M13mp8 previously digested with EcoRI and SmaI. By site directed mutagenesis a Xhol-site was inserted just in front of the ATG codon:

** *

5' ... TAAAATAAAAGCTCGAGATGAGATTT ... 3'.

Positions where bases have been exchanged by site directed mutagenesis, are indicated by asterisks.
Fig.4
pMaGAM-B/S. The BgIII fragment containing the glucoamylase gene and its 5' and 3' flanking region was subcloned into the BamHI site of vector pMAC5-8 (Kramer et al., 1984). By site directed mutagenesis using the gapped duplex DNA methode (Kramer et al., 1984) a BamHI site and a SalI site were inserted just in front of the translation initiation codon ATG:

* ***** *

5'... CTCATGACTG TGTCGACGGATCCAAGATGATTT...3'.

5 Positions affected by the site specific mutagenesis are indicated by asterisks.

Fig. 5
YRp7 α GAM. The GAM gene was isolated from pMaGAM-B/S as a Sall fragment and ligated into the XhoI site of M13 α Xho (Fig.3). The BgIII fragment harbouring the α -amylase promoter GAM gene fusion was ligated into the single BamHI site of the *S. cerevisiae* vector YRp7 resulting in plasmid YRp7 α GAM.10 Fig.6
Construction of *S. occidentalis* vectors for expression of foreign genes using the promoter and the signal sequence of the glucoamylase gene.

15 a: Insertion of the 4.0 kb EcoRI-PvuII fragment from pBRSwARSGAM, containing the complete GAM promoter and the first 208 bp of the GAM encoding sequence into pRR322, cleaved with EcoRI/PvuII, resulting in vector pBRGAM.

b: Insertion of the 3.4 kb PvuII fragment from plasmid pCT803, containing the structural gene coding for cellulase from base pair position +112, into the PvuII site of plasmid pBRGAM, resulting in vector pBRGC1.

20 c: Insertion of the 5.5 kb BgIII-PstI fragment from plasmid pBRGC1 into BamHI/PstI cut pCJD5-1. The resulting plasmid, pMPGC1-1 contains 321 bp of the GAM 5' non coding region and the first 208 bp of the GAM coding region fused to position +112 of the celD gene.

d: Insertion of the 6.5 kb BamHI-PstI fragment from plasmid pBRGC1 into BamHI/PstI cut pCJD5-1. The resulting plasmid, pMPGC1-2 contains 1.3 kb of the GAM 5' non coding region and the first 208 bp of the GAM coding region fused to position +112 of the celD.

25 Fig. 7
Promoter studies using the qualitative endoglucanase D activity assay (congo red staining) with different transformants:
NGA58: pMPAGC, NGA58: pMPPGC, NGA58: pMPGC1-1, NGA58: pMPGC1-230 a: Growth on 2% maltose
b: Growth on 4% glucoseFig.8
Construction of a *S. occidentalis* vector for expression of foreign genes under control of the ADH1 promoter.
a: Insertion of the 1.45 kb BamHI-Sall ADH1 promoter fragment of plasmid pFM2-1 into pCJD5-1, previously cut with BamHI/Sall resulting in plasmid pMPADH1.
35 b: Insertion of the 3.2 kb Sall fragment of pJDcg-15 into the Sall site of pMPADH1, resulting in pMPAG.
c: Exchange of the GAM promoter in plasmid pMPGC1-2 (Fig. 6d) for the ADH1 promoter from plasmid pMPAG by in vivo recombination, resulting in plasmid pMPAGC.40 Fig. 9
40 Construction of a *S. occidentalis* vector for the expression of foreign genes under control of the PDC promoter.
a: Insertion of the 6.0 kb SphI-Aval fragment from plasmid pJDcg-15, containing the PDC promoter fused to the GAM gene into YRpJD2 previously cleaved with SphI-Aval, resulting in plasmid pMPPG.
b: Exchange of the GAM promoter in plasmid pMPGC1-2 (Fig. 6d) for the PDC promoter from 45 plasmid pMPPG by in vivo recombination, resulting in plasmid pMPPGC.Fig. 10
Construction of an autonomous replicating vector (containing SwARS2) for transformation of *S. occidentalis*. The 3.2 kb TRP5 BamHI fragment from plasmid YRpJD2 was ligated with the 5.5 kb BamHI-BgIII fragment of plasmid pBRSwARSGAM (Fig. 6a), resulting in pMPTS2-A and pMPTS2-B.

50 ABBREVIATIONS USED THROUGHOUT THE APPLICATION

55 **Restriction endonucleases :**Aa-AatII
A - Asp718

B - BamHI
 Bg - BgIII
 C - ClaI
 E - EcoRI
 5 EV - EcoRV
 H - HindIII
 P - PvuII
 Ps - PstI
 S - SalI
 10 Sc - Scal
 Sm - SmaI
 Sp - SphI
 X - XbaI

Throughout this application various publications are referenced by mention of the first author and the
 15 year of publication within parentheses. Full citation of the references may be found at the end of the
 Specification as an annex, listed according to their alphabetical order immediately following the specifica-
 tion. The disclosures of these publications in their entireties are hereby incorporated by reference into this
 application in order to more fully describe the state of the art as known to those skilled therein as of the
 date of the invention described and claimed herein.

20 The yeast molecule according to the present invention is characterized by containing an expression
 cassette which in turn is composed of three to four different components, being arranged in operable
 linkage in order to ensure efficient expression of a desired polypeptide, optionally followed by secretion of
 same. These components are defined to be the following :

25 a) A first DNA sequence serving as a regulon. The term "a regulon" as used throughout the
 application comprises any cis-acting DNA sequence involved in the regulation of expression of a given
 structural gene. The term thus embraces sequences preceding a given coding sequence, for example, a
 promoter and sequences which are recognised by transcription factors or other proteins involved in
 transcription regulation. The "first DNA sequence" is thus not limited to sequences corresponding to known
 30 promoters. Also included are DNA sequences corresponding to natural occurring regulons, which have
 undergone insertion, deletion or substitution events, but still retain their activity as regulon. Further the term
 "regulon" includes DNA sequences involved in regulation of transcription, which have been synthesised
 chemically or obtained by methods of gene technology.

35 b) An optional second DNA sequence coding for a signal peptide. The meaning of this second DNA
 sequence includes various DNA sequences, directing the secretion of any polypeptide containing the thus
 encoded peptide at its N-terminal end.

40 c) A third DNA sequence coding for a foreign protein. This coding sequence coding for the signal
 peptide in order to retain a proper reading frame for the third DNA sequence. The fusion thus requires
 application of methods or recombination DNA technology known to every person skilled in the art. For
 methods enabling to design proper fusions it is in general referred to Maniatis et al. (1982), wherein
 45 appropriate methods, e.g. various ways to treat restriction fragment ends, cloning of linker molecules and
 similar procedures are described in great detail.

The foreign protein coded for by the third DNA sequences may be any protein which to express there
 exists some need. Since the secretion system is capable of secretion not only of small molecules, but also
 of large proteins, expression of any desired protein should be possible. For examples it is referred to the
 50 discussion below.

55 d) Optionally a fourth DNA sequence serving as a terminator. As the terminator any sequence may
 be used, which efficiently terminates transcription in the respective host organism. Terminators may be for
 example sequences prone to formation of hairpin structures and/or polyadenylation sites. In a preferred
 embodiment the terminator is derived from the same gene as the second DNA sequence, i.e. the regulon, is
 derived from.

The four components are combined with each other using known techniques. Optionally the proper
 fusion will be verified for example by sequencing the boundaries.

The yeast cell according to the present invention contains at least one expression cassette comprising
 some or all of the above mentioned DNA sequences. It provides a convenient tool for the industrial
 55 expression of any desired protein.

In a preferred embodiment said first and/or second and/or fourth DNA sequence are derived from a
 gene coding for an amylolytic enzyme, for example α -amylase or glucoamylase, which enzymes in most
 organisms are induced upon contact with starch. The DNA sequences may be derived from a yeast of a

member of one of the genera Debariomyces, Saccharomyces, Lipomyces, Pichia, or Saccharomyces, but yeast of the genus Schwanniomyces are preferred.

5 Yeast species of the genus Schwanniomyces have an efficient expression and secretion apparatus, enabling them to grow on starch as a sole carbon source. Starch is hydrolysed to glucose by two amylolytic enzymes, secreted into the supernatant, namely α -amylase (α -1,4-glucan 4-glucanohydrolase E.C.3.2.1.1) and glucoamylase (syn. amyloglucosidase) (α -1,4-glucan glucohydrolase, E.C.3.2.1.3.; with debranching activity E.C.3.2.1.9.). The amylolytic enzymes of Schwanniomyces occidentalis, formerly also called Gyngschwanniomyces castellii and Schwanniomyces alluvius (Price et al., 1978), are well documented (Oteng-Gyeng et al., 1981; Sills et al., 1982, 1984a, 1984b; Wilson et al., 1982).

10 The genes for α -amylase and glucoamylase from a yeast of the genus Schwanniomyces have been disclosed by the present inventors in EP 87 110 370.1. However, at that time the DNA or peptide sequences directing secretion have not been identified. Furthermore their use for the establishment of an efficient expression and secretion system in yeast was not known then.

15 It is preferred, to use as a first DNA sequence part or all of a 1.8 kb BgIII-Xhol fragment which precedes the structural gene coding for the Schwanniomyces α -amylase, or part or all of a 1.3 kb BamHI-PvuII fragment preceding the structural gene coding for the Schwanniomyces occidentalis glucoamylase. Further if only smaller fragments are to be used, it is preferred, to use a DNA sequence comprising part or all of the DNA sequence corresponding to bases -1 to -540 of the region preceding the structural gene coding for α -amylase or part or all of the DNA sequence corresponding to bases -1 to -320 of the region preceding the structural gene coding for glucoamylase. The respective DNA fragments are fully active as promoters, inducible by starch, dextrin, maltose.

20 It is obvious, that replacement of bases or base pairs, not involved in the transcriptional regulation, still are within the spirits of the present invention. Thus natural mutations or synthetic equivalents also are embraced by the present invention.

25 In yeast cells of the genus Schwanniomyces it is also possible to use promoters derived from yeast genes obtained from many other yeast genera. Preferred are regulons involved in the expression of ADH1, ADH2, PDC1, GAL1/10, PGK and GAPDH, or LAC4 which are obtained from Saccharomyces cerevisiae, Kluyveromyces Lactis or Schwanniomyces occidentalis.

Furthermore, it is possible to use viral regulons and/or terminators, for example regulons and terminators obtained from E.coli T-phages. The phages possess extraordinarily strong functional sequences for performance of their lytic functions. In a preferred embodiment sequences of phage T7 are used, which are regulated by expression of T7 RNA polymerase.

30 In order to increase the over all expression level and furthermore, in order to facilitate purification of produced polypeptides, it is most desirable, that the polypeptides are secreted into the culture medium. The present invention provides signal sequences having the invaluable advantage of being derived from genes of amylolytic and thus inducible enzymes, which upon induction have to be secreted very efficiently. A further advantage is, that induction of amylolytic enzymes does not require complicated media, lacking for example a specific amino acid or phosphate, but only require the culture to be transferred to a starch containing medium. Signal sequences, which are preferably used for secretion of large proteins include part 35 or all of at least one of the following peptides:

- 40 a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaLeuAlaPheSerArgLeuValSerAla; LeuAlaPheSer-
- b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaLeuAlaPheSer-
- ArgLeuValSerAlaGlnProllellePheAspMetArg;
- c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAla;
- d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAlaAlaGlnAlaAlaProAla;
- e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAla-

45 IleGlnAlaAlaProAlaSerSerIleGlySerSerAlaSerAla

In a further embodiment the above mentioned peptides are coded for by DNA sequences, which correspond to all or part of the following DNA sequences:

- 50 a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT AGCATTCTCAAGATT- GGTATCTGCT;
- b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT TAGCATTCTCAAGATT- GGTATCTGCTCAACCGATTATTTTGACATGAGA;
- c) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA GTTAGTGCT;
- d) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA GTTAGTGCTATCCAAG- CAGCCCCCTGCC;
- e) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA GTTAGTGCTATCCAAG- CAGCCCCCTGCCCTTCGATTGGATCTAGTGCTCA GCA.

It will be understood, that the replacement of amino acids by residues having comparable properties and use of alternative codons in compliance with the possibilities provided by the genetic code will not have any impact on the yeast cell according to the present invention.

The polypeptide to be expressed in a yeast cell according to the present invention is coded for by a third DNA sequence. Said third DNA sequence may comprise natural or synthetic DNA, even DNA containing intervening sequences. Examples for proteins being expressed by the yeast cell according to the present invention comprise proteins of enormous importance for scientific or medical purposes, for example cellulase, interleukine, interferon, insulin like growth factor, lymphokine, human growth factor, nerve growth factor, aprotinin, insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines or human granulocyte/macrophage colony stimulating factor.

The yeast cell according to the present invention may optionally comprise a terminator as component of said expression cassette. Preferred terminators are derived from genes of one of the following genera : Saccharomyces, Pichia, Hansenula, but preferably Schwanniomyces. Examples of preferred terminators are those derived from Schwanniomyces α -amylase gene or Schwanniomyces glucoamylase gene, part or all of which may be used. Such terminators are comprised by nucleotides 1537 to 1740 of the Schwanniomyces α -amylase gene or nucleotides 2875 to 3320 of the Schwanniomyces glucoamylase gene.

It is obvious, that in case of any of the DNA sequences mentioned modifications, which do not impair the biological function of the respective DNA fragment still are covered by the present application. There are numerous possibilities to change the DNA sequences mentioned above slightly without impairing the biological function of said fragment gene.

The yeast cell according to the present invention, carrying an expression cassette composed of the above discussed components, may be carried by a circular or linear vector, which optionally comprises one or more selective marker genes. The present inventors demonstrated, that a lot of marker genes so far used for transformation of Saccharomyces, for example TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, which may be obtained from Saccharomyces, Hansenula, Pichia or Schwanniomyces, are functional in Schwanniomyces occidentalis. Furthermore it is also possible to use the amylolytic properties of α -amylase or glucoamylase in order to select transformed organisms. The corresponding genes have been isolated and characterized. AMY1 and GAM1 are preferably obtained from Schwanniomyces.

The yeast cell according to the present invention may contain the expression cassette as an insert into one of the endogenous yeast chromosomes. Furthermore, there is however the possibility, that a DNA sequence capable of controlling autonomous replication and even also stable maintenance of the vector in the host organism, is introduced into the yeast cell. Said DNA sequences are the so called ARS sequences, capable of controlling autonomous replication and stable maintenance in the respective host organism. Said DNA sequences are preferably derived from Saccharomyces, Pichia Hansenula or Kluyveromyces, most preferred from Schwanniomyces.

In order to ensure effective replication in the prospective host organism, it is preferred to use an ARS, derived from the respective host organism.

The present invention provides one new SwARS sequence. The DNA sequence SwARS2 is capable of controlling autonomous replication and stable maintenance of the vector in a Schwanniomyces yeast cell. The detailed sequence of SwARS has not yet been determined. SwARS2 is localised in the Schwanniomyces genome in the 3' region of the glucoamylase gene, as shown for example in Fig. 10. SwARS2 is easy to identify and efficiently allows autonomous replication of any plasmid, exhibiting this sequence. The so far known Schwanniomyces autonomously replicating sequences, SwARS1 and SwARS2 are comprised on a Clal-BamHI fragment (SwARS1) adjacent to the EcoRI-fragment shown in Fig 2a (see Fig. 10) and/or an EcoRI-BglII fragment contained in the EcoRI fragment shown in Fig 2a (SwARS2).

For the SwARS sequences it is also not necessary, in any case to use the sequence as it occurs naturally. DNA sequences, that have been obtained by any kind of modifying treatment, for example by deletion, insertion or substitution of one or more nucleotides or base pairs, are still within the scope of the present invention, provided the modifications retain or improve the biological function of the yeast cell according to the present invention.

In a preferred embodiment the yeast cell according to the present invention comprises a DNA sequence, which is homologous to genomic Schwanniomyces DNA. Provision of structures of homologous DNA enable the DNA fragment flanked by such homologous sequences to be inserted by homologous recombination. It is thus preferred to have homologous sequences, which do not represent any essential gene of the Schwanniomyces genomic DNA, but rather any gene, whose expression product may be supplemented by additives to the respective medium or were lack of the respective expression product does not have any serious impact on the transformed yeast cell.

If no SwARS sequence or any equivalent sequence, for example originating from any other yeast

species, is available, the expression cassette can be used without an autonomously replicating agent. In this case the expression cassette tends to integrate into the genome, for example via homologous recombination. Preferably there are several copies inserted into one or more chromosomes.

5 In this case it is preferred, to have multiple copies of the expression cassette inserted, in order to ensure simultaneous transcription and translation of several copies.

When the yeast cell according to the present invention is derived from a yeast of the genus Schwanniomyces the polypeptide products observed following expression of the respective foreign gene in said yeast cells exhibit glycosylation patterns, which closely resemble or even match those of the desired 10 eucaryotic protein. Yeast cells of the genus Schwanniomyces are thus preferred for expression of eucaryotic proteins.

15 The above discussed yeast cell according to the present invention can be used in a process for the production of a polypeptide, wherein a yeast host organism is cultivated under suitable conditions and the polypeptide is recovered in a manner known per se. The conditions of cultivation depend on the respective host organism used. Examples for cultivation conditions for yeast of the genus Schwanniomyces are provided in the examples below. The polypeptide, produced by the yeast according to the present 20 invention, is recovered by conventional purification methods, which depend on the properties of the expressed protein, the availability of specific absorbing agents and so on.

25 In a preferred embodiment yeast cells of the species Schwanniomyces occidentalis are used as a host organism and are cultivated on starch,dextrin, maltose and/or plant biomass as the respective carbon source. As commonly known, these carbon sources are available at low costs.

Furthermore, there is a possibility to use cells of the genus Schwanniomyces for the production of single cell protein. The single cell protein may then be subjected to any further use as it is required.

30 In a preferred embodiment the polypeptide produced in a process according to the present invention is an amylolytic enzyme, for example an α -amylase or glucoamylase. These enzymes are commercially used to debranch and degrade starch, which thereupon may be further processed, for example in brewery or 25 baking processes. There is a high need for the provision of starch degrading enzymes.

35 In a further embodiment of the process according to the present invention production of a polypeptide may be achieved by providing expression cassettes as discussed above containing a yeast α -amylase or glucoamylase signal sequence and inserting the respective expression cassette into a yeast of one of the genera Saccharomyces, Kluyveromyces, Hansenula, Pichia or Schizosaccharomyces. Production and secretion 40 of foreign proteins in the respective yeast host organism is obtained by applying the appropriate inductive measure, whereupon the produced protein is secreted due to the presence of the yeast α -amylase or glucoamylase signal sequence. Since it is known from experiments performed by the present inventor, that these signal sequences are recognised in virtually any yeast genus, the α -amylase or glucoamylase 45 signal sequences may be applied in many different yeast with success. Also in this case, it is preferred, to use the α -amylase or glucoamylase signal sequences derived from the respective genes of Schwanniomyces occidentalis.

45 In summary the yeast cells according to the present invention are well suited for transformation and expression of any desired foreign gene. As already outlined above, expression of prokaryotic as well as eucaryotic genes is possible. However, for expression of eucaryotic proteins yeast cells of the genus Schwanniomyces containing some or all of the DNA sequences mentioned above are the most preferred yeast cells.

45 CHARACTERISATION OF THE α -AMYLASE AND GLUCOAMYLASE GENE EXPRESSION IN DIFFERENT YEAST GENERA :

50 The N-termini of the deduced amino acid sequences of the α -amylase gene (Fig.1B) and glucoamylase gene (Fig.2B) exhibit the features of typical leader sequences of secretory proteins (Von Heijne 1983, Perlman and Halvorsson 1983). The N-terminal amino acids of mature α -amylase secreted by S. occidentalis are determined to be AspValSer. This indicates processing between Arg33 and Asp31 which is unusual and not according to the Von Heijne's model of site specificity for signal peptidase cleavage. This putative cleavage site was confirmed also for α -amylase expressed in S. cerevisiae, Pichia stipidis, K. lactis, S. pombe and H. polymorpha. However, in these yeast species processing is also found to occur between Ala 55 25 and Gln 26, resulting in the N-terminal sequence GlnProlle.

55 α -amylase isolated from a culture of S. occidentalis migrates as a single band of 55 kDa on SDS polyacrylamide gels: following Endoglycosidase H treatment and subsequent gel electrophoreses a single band of 54 kDa is observed, indicating that only one of the two potential N-glycosylation sites of the protein

is used for glycosylation. It is noteworthy that high mannose glycosylation, as very often found by proteins secreted in S. cerevisiae is not accompanying secretion in S. occidentalis.

5 IDENTIFICATION OF THE TRANSCRIPTION INITIATION SITES AND PROMOTER FUNCTION OF THE α -AMYLASE AND GLUCOAMYLASE GENE.

10 By S1 mapping of polyA-mRNA (Sharp et al., 1980) the transcription initiation site of the α -amylase gene could be localised to be in the region of -42 to -30 (indicated by an interrupted line above the relevant sequence in Fig.1B) and the transcription initiation site of the glucoamylase gene was determined to be in the region -7 to -10 (underlined in Fig.2B).

CONSTRUCTION OF AN α -AMYLASE - GLUCOAMYLASE HYBRID GENE

15 In order to facilitate use of the α -amylase promoter and glucoamylase structural gene in further constructions additional restriction sites were inserted in front of the translation initiation codon of both genes using oligonucleotide directed mutagenesis as described in the AMERSHAM protocol. Correct mutagenesis was verified by sequencing according to Sanger (1977). The boundaries between the original 20 Schwanniomyces genes and the inserted polylinker are shown below :
 α -amylase :

** *

25 5' ... TAAAATAAAAGCTCGAGATGAGATTI ... 3' .
XbaI

Glucoamylase:

30 * ***** * *
5' ... CTCATGACTGTGTCGACGGATCCAAGATGATTTT ... 3' .
SalI BamHII

The manipulations resulted in plasmids M13 α Xho (Fig.3) and pMaGAM-B/S (Fig.4).

40 Thus, both the α -amylase promoter and the structural glucoamylase gene can be isolated from plasmid M13 α Xho and pMaGAM-B/S as suitable DNA fragments. The GAM gene was isolated from pMaGAM B/S as a 3.2 kb Sall fragment and ligated into the XhoI site of M13 α Xho. The BgIII fragment harbouring the α -amylase promoter/GAM gene fusion was ligated into the single BamHI site of the S. cerevisiae vector YRp7 45 (Struhl et al., 1979) resulting in plasmid YRp7 α GAM (Fig. 5). S. cerevisiae strain YNN27 was transformed with this plasmid selecting for TRP prototrophy. Transformants secreted active glucoamylase into the culture supernatant (50 mU/ml) indicating that the α -amylase promoter fragment can direct expression and secretion of foreign proteins.

50 SECRESSION OF α -AMYLASE AND GLUCOAMYLASE IN DIFFERENT YEASTS USING SUITABLE PROMOTERS.

55 Expression of the α -amylase and glucoamylase gene and secretion of the gene products can be detected in K.lactis under control of the promoter of the α -galactosidase (LAC4) gene (Breunig et al., 1984). Expression of the α -amylase and glucoamylase gene and secretion of the both proteins can also be detected in S. pombe under control of the ADH1 promoter (Russel et al., 1983) or of the GAL1/10 promoters (Johnston and Davis 1984). After fusion with the GAL1/10 promoter (Johnston and Davis 1984), DHAS promoter (Janowicz et al., 1985), MOX promoter (Ledeboer et al., 1985) or FMDH promoter (EP 87 110

417.0) expression of the α -amylase and glucoamylase gene and secretion of both proteins can be detected in Hansenula polymorpha.

5 EXPRESSION AND SECRETION STUDIES IN S. OCCIDENTALIS USING THE CELLULASE EN-
DOGLUCANASE D (EGD) OF CLOSTRIDIUM THERMOCELLUM AS A MODEL FOR FOREIGN GENE
EXPRESSION.

As a first example for the expression of a foreign gene in S. occidentalis, the celD gene from C. thermocellum (Millet et al., 1985, Joliff et al., 1986b), which codes for a thermostable cellulase, namely an endoglucanase (EGD) was used. Advantages in using this system are :

- specific reaction of EGD, which can easily be monitored
- a rapid qualitative assay for colonies with EGD activity (see material and methods)
- availability of antibodies against EGD for specific detection of translation products.

15

CHARACTERISATION OF THE GLUCOAMYLASE PROMOTER AFTER FUSION WITH THE celD GENE

Plasmids were constructed containing a replicon from S. occidentalis (SwARS1) (EP 87110370.1) the TRP5 gene from S. cerevisiae (Zalkin and Yanowski, 1982) as a selective marker and, in addition, a GAM/celD gene fusion under the control of different promoters. In the first step the 4.0 kb EcoRI-PvuII fragment from plasmid pBRSwARSGAM (Fig. 6a) was isolated and inserted into the 2296 bp pBR322 EcoRI-PvuII fragment, containing the amp gene and the bacterial origin, resulting in plasmid pBRPGAM (Fig. 6a). In addition to the pBR322 sequence this plasmid carries 3.6 kb from the 5' non coding region of the GAM gene and the first 208 bp coding for the N-terminal part (including the signal sequence) of the glucoamylase. A 3.4 kb PvuII fragment from plasmid pCT603 (Joliff et al., 1986) containing the coding region of the celD gene missing the 5' 111 bp was inserted into the PvuII single site of pBRPGAM resulting in pBRGC1 (Fig. 6b). E. coli transformed with this plasmid results in weak Carboxymethylcellulase (CMCase) activity, analysed by congo red staining (see material and methods) indicating that the glucoamylase promoter is slightly expressed in E. coli. For the construction of a S. occidentalis expression vector the plasmid pCJD5-1 (EP 87110370.1) was cleaved with BamHI-PstI and ligated with the 5.5 kb BglII-PstI fragment or 6.5 kb BamHI-PstI fragment from pBRGC1, resulting in pMPGC1-1 (Fig. 6c) and pMPGC1-2 (Fig. 6d), respectively. Both plasmids share an in frame fusion of GAM/celD but differ in the length of the 5' upstream region of the GAM gene. S. occidentalis strain NGA58 (trp5 ade) isolated after UV mutagenesis of strain NGA23 was transformed with these plasmids and transformants selected for tryptophan prototrophy and analyzed by means of congo red assay for cellulase activity. Transformants produced active EGD when grown under inducing conditions in 2% maltose (Fig. 7a). By analysis with the congo red assay no activity could be detected under repressed conditions in 4% glucose (Fig. 7b). NGA58:pMPGC1-1 and NGA58:pMPGC1-2 transformants secrete active EGD into the culture supernatant when grown in maltose under induced conditions. The intact or complete glucoamylase promoter resides on plasmid pMPGC1-2 as shown by 20 times higher activity compared to the shorter promoter fragment in gene fusion of pMPGC1-1 (table 1). However the 0.3 kb promoter fragment still harbours regulatory units which lead to catabolite repression by glucose (Fig. 7b). Comparison of the N-terminal GAM sequence to the signal peptidase cleavage sites of several other secretory proteins (Von Heijne 1983) reveals three potential signal peptidase cleavage sites after the amino acid residues Ala 19, Ala 25 and Ala 34. For determination of the N-terminal amino acid of the mature glucoamylase the protein was isolated for N-terminal amino acid sequencing. However the N-terminus is blocked and therefore could not be sequenced. To confirm the correct cleavage site of the GAM signal sequence in S. occidentalis, the regions between the potential signal peptidase cleavage sites and the PvuII gene fusion site are deleted in the GAM/celD fusion by in vitro mutagenesis to study the influence on secretion of active EGD.

STUDIES OF THE USE OF SUITABLE S. CEREVIAE PROMOTERS IN THE EXPRESSION OF FOREIGN GENES IN S. OCCIDENTALIS.

55 The ADH1 promoter (Hitzeman et al., 1981) was isolated as a 1.45 kb BamHI-SalI fragment from plasmid pFM2-1 (Müller et al., 1987) and ligated into plasmid pCJD5-1 (EP 87 110370.1) after cleavage with BamHI and SalI resulting in pMPADH1 (Fig. 8a). The GAM gene was excised as a SalI fragment from

pJDcg-15 (see below) and inserted into the single Sall site of pMPADH1 (Fig. 8b) resulting in plasmid pMPAG (Fig. 8b). By the means of in vivo recombination (Ma et al., 1987) the GAM promoter from pMPGC1-2 (Fig 6d.) was replaced by the ADH1 promoter of pMPAG leading to plasmid pMPAGC (Fig. 8c). Plasmid pMPGC1-2 was linearized with BamHI; the unique BamHI site of pMPGC1-2 is localized between 5 the TRP5 coding sequence and the glucoamylase promoter. The 2.75 kb Clal fragment of pMPAG containing the ADH1 promoter shares homologies to the TRP5 gene as well as to the GAM coding region up to position +208. For in vivo recombination NGA58 was transformed with BamHI linearized pMPGC1-2 together with approx. 10 fold excess of Clal fragment using the transformation protocol of Ito et al., 1983.

10 Recombinants were screened for their ability to hydrolyze CMC on 4% glucose. Under these conditions celD expression is repressed in transformants harbouring recircularized pMPGC1-2. From 146 transformants 7 colonies expressing cellulase in the presence of 4% glucose were isolated. NGA58 containing the "in vivo" constructed plasmid pMPAGC expresses active EGD constitutively (see Fig. 7). This result demonstrates that the newly constructed plasmid pMPADH1 serves as a vector for foreign gene expression.

15 In another experiment a recombinant plasmid was constructed by in vivo recombination to study the expression of the GAM/celD gene fusion under control of the PDC promoter (Das and Hollenberg, 1982). Therefore the 6.0 kb SphI-Aval fragment from plasmid pJDcg-15 (EP 87 110370.1) carrying the glucoamylase gene fused to the PDC promoter was subcloned into the respective restriction sites of YRpJD2 (EP 87 110370.1) resulting in plasmid pMPPG (Fig. 9a).

20 For in vivo recombination pMPPG was cleaved with Clal and used together with BamHI linearized pMPGC1-2 for transformation of NGA58. Homologous recombination results in plasmid pMPPGC (Fig. 9b) harbouring a GAM-celD fusion under control of the PDC promoter. This construction leads to an expression in Schwanniomyces occidentalis NGA58 inducible by glucose and partly repressible by maltose (Table 1). Transformants were analysed for expression of active EGD using the qualitative congo red assay (Fig. 7). The results indicate that the 0.3 kb GAM promoter fragment leads to regulated expression, however less 25 efficient than the 1.3 kb fragment (for comparison see table 1 and Fig. 2A). Expression under the ADH1 promoter leads to efficient constitutive expression. The expression level is comparable with that of induced expression directed by the 1.3 kb GAM promoter.

25 The fact that plasmids can be constructed by in vivo recombination is a proof for homologous recombination and hence offers the possibility for stable site directed integration into the genome of S. occidentalis, a prerequisite for the industrial production of foreign proteins.

30 The expression of the α -amylase gene and glucoamylase gene and secretion of the both gene products in S. occidentalis could also be achieved under control of the phosphoglycerate kinase (PGK) promoter (Dobson et al., 1982), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter (Holland and Holland 1979), copperchelatin (CUP1) promoter (Karin et al., 1984, Butt et al., 1984) and alcoholdehydrogenase 2 35 (ADR2 or ADH2) promoter (Russell et al., 1983, Beier et al., 1985) or galactokinase (GAL1/10) promoter (Johnston and Davis, 1984).

Construction of pJDcg-15 :

40 The GAM gene was isolated as a 3.2 kb BamHI-Sall DNA fragment from plasmid pMaGAM B/S (Fig.4) and inserted into the respective restriction sites of plasmid pBM272 Johnston & Davis, 1984). The PDC promoter was isolated from plasmid pCP202 (Kellermann & Hollenberg, 1988) as a Sall-SphI fragment and ligated into the respective single sites of vector YRp7 (Struhl et al., 1979). To be able to isolate the PDC1 45 promoter as a BamHI fragment the 3.2 kb Sall fragment of PMaGAM B/S containing the GAM gene was fused to the PDC promoter via the single Sall site. The resulting BamHI fragment containing the PDC-promoter was ligated into the single BamHI site of pBM272 containing the GAM gene. The final plasmid is called pJDcg-15 (Fig. 8b).

50 EXPRESSION OF HUMAN GRANULOCYTE/MACROPHAGE COLONY STIMULATING FACTOR (hGM-CSF)

55 The gene coding for hGM-CSF (Cantrell et al., 1985) was obtained from British Biotechnology Ltd. as a HindIII-EcoRI fragment. Position +52 of hGM-CSF gene was fused to position + 102 of the GAM signal sequence under control of the ADH1 or GAM promoter. This fusion results in a hybrid protein in which the endopeptidase cleavage site of the GAM signal sequence is linked with a Met included in the synthetic gene offered by British Biotechnology Ltd., followed by the first amino acid of mature hGM-CSF (Ala 18).

This fusion was inserted into a S. occidentalis expression vector and transformed into NGA58. Western blot analysis revealed secretion of hGM-CSF protein.

5 EXPRESSION OF INTERLEUKINE-2

The gene coding for IL-2 (Taniguchi et al., 1983) was obtained from British Biotechnology Ltd., as a EcoRI-HindIII fragment. Position +61 of IL-2 gene was fused to position + 102 of the GAM signal sequence under control of the ADH1 or GAM promoter. This fusion results in a hybrid protein in which the 10 endopeptidase cleavage site of the GAM signal sequence is linked with a Met included in the synthetic gene offered by British Biotechnology Ltd., followed by the first amino acid of mature IL-2 (Ala 21). This fusion was inserted into a S. occidentalis expression vector and transformed into NGA58. Western blot analysis revealed secretion of IL-2 protein.

15

ISOLATION OF AN ADDITIONAL SwARS ELEMENT AND A NEW SELECTABLE MARKER

A new SwARS sequence (SwARS2), localized downstream of the coding region of the GAM gene, could be identified. To analyze the function of SwARS2 a vector was constructed containing the Trp5 gene and 20 SwARS2 in pBR322, resulting in pMPTS2-A and pMPTS2-B, respectively (Fig. 10). For the construction of pMPTS2-A and pMPTS2-B the 5.5 kb BamHI - BgIII fragment of pBRSwARSGAM, constructed by insertion of the EcoRI-fragment shown in Fig.2a into the single EcoRI site of plasmid pBRSwARS1, (disclosed in EP 87 110370.1), was isolated and ligated with 3.2 kb BamHI TRP5 fragment obtained from YRpJD2. High frequency transformation of S. occidentalis is obtained using this plasmid.

25 As an additional selectable marker the LEU2 gene of S. occidentalis was isolated. The LEU2 gene, coding for isopropyl malate dehydrogenase, was cloned by functional complementation of the mutant AH22 (his4leu2) (Hinnen et al., 1978) of S. cerevisiae by transformation with a cosmid genomic library from S. occidentalis.

30 The LEU2 gene was found to reside on a 9 kb EcoRI fragment of Schwanniomyces genomic DNA; it's function could be abolished by insertion of the TRP5 gene of S. cerevisiae. The disrupted LEU2 gene was integrated into the genome of NGA58 by homologous recombination resulting in a new strain NGA581 (TRP5, LEU2, and ADE). An analogous experiment with the cloned HIS4 gene of S. occidentalis (EP 87 110 370.1) led to a strain NGA58h (TRP5, HIS4, ADE).

35

IMPROVED EXPRESSION OF HETEROLOGOUS GENES USING THE T7 RNA POLYMERASE AND T7 PROMOTER AND TERMINATOR FOR THE EXPRESSION OF FOREIGN GENES IN S. OCCIDENTALIS

40 T7 RNA polymerase (Fuerst et al., 1986) can be expressed in S. occidentalis after fusion of the structural gene (Dunn and Studier, 1984) using the BamHI restriction site near the translation initiation codon with suitable promoters as α -amylase, GAM (0.3 kb fragment) or PDC and homologous integration into the genome via any homologous cloned gene, e.g. HIS4, LEU2, AMY1 and GAM. A suitable fragment, containing the fusion flanked by homologous cloned sequences can be isolated and used together with pCJD5-1 (Fig. 6c) for co-transformation of NGA58, selecting on YNB containing 0.5% casaminoacids.

45 Colonies containing the integrated T7 RNA polymerase gene can be identified by Southern analysis after isolation of transformants mutated in the target gene.

50 For the expression of foreign genes a S. occidentalis expression vector was constructed containing SwARS2, the ColE replication origin of E. coli, a suitable selective marker for S. occidentalis (e.g. TRP5, HIS4, LEU2) and a cassette containing the T7 polymerase promoter and terminator sequences isolated from plasmid pAR2529 (Fuerst et al., 1986), separated by the polylinker sequence from M13mp19. To study the efficiency of the T7 system we expressed a GAM/celD fusion (GAM signal sequence and structural celD gene obtained as a BamHI-Asp 718 fragment from pMPGC1-2) under control of the T7 promoter by insertion into the BamHI-Asp718 site of the polylinker sequence. Using this system active cellulase is secreted into the culture medium.

55

MATERIALS AND METHODS

1. The microorganisms used are as follows :

Schwanniomyces occidentalis NGA23 : (DSM 3792)Schwanniomyces occidentalis NGA585 Kluyveromyces Lactis SD11 : (DSM 3795)Schizosaccharomyces pombe LEU1-32, HIS5-303 (DSM 3796)Saccharomyces cerevisiae YNN27 (URA3-52, TRP1-298, GAL2)10 The Escherichia coli strain used is as follows E. coli HB101 F-, hsd S20 (r_B- m_B-), recA13, ara-14, proA2,
lacY1, galK2, rpsL20 (Sm^r), xyl-5, mtl-1, supE44)

2. Culture media :

15 Schwanniomyces occidentalis strain NGA58 (isolated by UV mutagenesis of strain NGA23) was grown at 30 °C in YNB (0.67% yeast nitrogen base without amino acids) supplemented with appropriate amino acids and with 2% of soluble starch when indicated.20 S. cerevisiae strain YNN27 (URA3-52, TRP1-289, GAL2) (Stinchomb et al 1980); K. lactis strain SD11 TRP1 (Das and Hollenberg, 1982); S. pombe strain LEU1-32, HIS5-303 (obtained from W.D. Heyer); These strains were grown under conditions as described above for Schwanniomyces occidentalis with the replacement of 2% soluble starch by 2% glucose, where indicated.25 The culture media for S. occidentalis and Hansenula polymorpha were buffered with 0.2 M NaPO₄ buffer pH 6.2; those for S. cerevisiae and K. lactis were buffered with 0.1 M citrate buffer pH 6 and those for S. pombe with 0.05 M acetate buffer pH 6.30 25 The pH optimum for the cellulase (Joliff et al., 1986a), α -amylase and glucoamylase (Wilson and Ingledew, 1982) is approx. pH 6.

E. coli HB101 (Bolivar et al., 1977) was grown in LB medium (Maniatis et al., 1982) with penicillin G (150 ug/ml).

3. Miscellaneous methods

35 Plasmid DNA was purified either by CsCl gradient centrifugation (Maniatis et al., 1982) or by rapid alkaline extraction of plasmid DNA described by Birnboim and Doly (1979).

40 Yeast crude extracts were prepared by the method of Schatz (1979). Yeast minilysates were prepared by the method of Sherman et al., (1983).

DNA fragments were isolated by the "low melting agarose" procedure as described by Gafner et al., (1983). Southern hybridization was carried out as described by Southern (1975).

45 Yeast transformation was performed according to Klebe et al. (1983) or Ito et al. (1983). The transformants or integrants were tested for expression and secretion of cellulase using either the qualitative congo red assay (Teather and Wood, 1982: halo formation on carboxymethylcellulose agar plates) or a quantitative cellulase assay according to Joliff et al. (1986a). One unit (U) is defined as the amount of enzyme that released 1 umol para-nitrophenol per min at 60 °C. Western analysis was performed as described by Towbin et al. (1979).

50 The transformants or integrants were tested for expression and secretion of α -amylase using either a standardized α -amylase enzyme test from Merck, W. Germany or a qualitative starch degradation test (halo formation on starch agar plates after staining with iodine), as follows : in this test the rate of formation of 2-chloro-4-nitrophenol is determined by photometry at 405 nm in 0.1 M potassium phosphate at 37 °C. The enzyme unit (U) is defined as the amount of enzyme catalysing the formation of 1 umol 2-chloro-4-nitrophenol per min at 37 °C.55 Glucoamylase activity was measured by a stop assay method: After incubation of the sample in 10% soluble starch in 0.05M KH₂PO₄-NaOH (pH 5.0) at 50 °C the amount of glucose produced is determined by the glucose dehydrogenase method (system glucose Merck). The quantity of NADH formed is proportional to the glucose concentration. The enzyme unit (U) is defined as the amount of enzyme catalysing the formation of 1 umol glucose/min at 50 °C.

TABLE 1

	Transformant	OD _{600 nm}	EDG-activity mU/ml
5	NGA58: pMPGC1-1 (0.3 kb promoter)	2.3 5.3	0.12 0.73
10	NGA58: pMPGC1-2 (1.3 kb GAM promoter)	2.0 7.9	0.53 1.83
15	NGA58: pMPAGC ADH1 promoter	3.0 5.9	0.52 1.07
20	NGA58: pMPPGC PDC1-promoter	2.2* 4.5* 2.9 7.2	0.40* 0.64* 0.19 0.31
25	Transformants were grown in YNB, 0.5% Casaminoacids, 2% maltose, 20 mg/l adenine and 0.2 M NaP buffer pH 6.2		

*: values for transformants grown in 4% glucose

30 LIST OF REFERENCES

Beier et al. (1985), Mol. Cell. Biol. 5, 1743-1749.
 Birnboim & Doly (1979), Nuc. Acids Res. 7, 1513-1523.
 Bitter et al. (1984), Proc. Natl. Acad. Sci. USA 81, 5330-5334.
 Bolivar et al. (1977), Gene 2, 75-93.
 Breunig (1984), NUCl. Acids Res. 12, 2327-2341.
 Butt et al. (1984), Proc. Natl. Acad. Sci. USA 81, 3332-3336.
 Cantrell et al. (1985), Proc. Natl. Acad. Sci. USA 82, 6250-6254.
 Das & Hollenberg (1982), Curr. Genet. 6, 123-128.
 Dobson et al. (1982), Nucleic Acids Research 10, 2625-2637.
 Dunn & Studier (1984), Proc. Natl. Acad. Sci. USA 81, 2035-2039.
 Emr et al. (1983), Proc. Natl. Acad. Sci. USA 80, 7080-7084.
 Fuerst et al. (1986), Proc. Natl. Acad. Sci. USA 83, 8122-8126.
 Gafner et al. (1983), Embo J. 2, 583-591.
 Hinnen et al. (1978), Proc. Natl. Acad. Sci. USA 75, 1929-1933.
 Hitzeman et al. (1981), Nature 294, 717-722.
 Holland & Holland (1979), J. Biol. Chem. 254, 5466-5474.
 Ito et al. (1983) J. Bacteriol. 153, 163-168.
 Janowicz et al. (1985), Nucleic Acids Res. 9, 3043-3062.
 Johnston & Davis (1984), Mol. Cell. Biol. 4, 1440-1448.
 50 Joliff et al. (1986a), Bio/Technology 4, 896-900.
 Joliff et al. (1986b), NUCl. Acids Res. 14, 8605-8613.
 Karin et al. (1984), Proc. Natl. Acad. Sci. USA 81, 337-341.
 Kellermann & Hollenberg (1988), Curr. Genetics 14, 337-344.
 55 Klebe et al. (1983), Gene 25, 333-341.
 Kramer et al. (1984), Nucl. Acids. Res. 12, 9441-9456.
 Kurjan & Herskowitz (1982), Cell 30, 933-943.
 Ledeboer et al. (1985), Nucl. Acids. Res. 9, 3063-3082.

Ma et al. (1987), *Gene* 8, 201-216.

Maniatis, Fritsch & Sambrook (1982), *Molecular cloning: A laboratory manual*, Cold Spring Harbour Laboratory press, New York.

Millet et al. (1985), *FEMS Microbiol. Lett.* 29, 145-149.

5 Müller et al. (1987), *Mol. Gen. Genetics* 207, 421-429.

Oteng-Gyang et al. (1981), *Zeitschrift allgemeine Mikrobiologie* 21, 537-544.

Perlman et al. (1983), *J. Mol.* 167, 391-409.

Price et al. (1978), *Microbiol. Rev.* 42, 161-193.

Russell et al. (1983), *J. Biol. Chem.* 258, 2674-2682.

10 Sanger et al. (1977), *Proc. Natl. Acad. Sci. USA* 74, 5 463-5 467.

Schatz (1979), *Meth. Enzymol.* LVI, 40-50.

Sharp et al. (1980), in: *Methods Enzymol.* 65,

Sherman et al. (1983), *Methods in yeast genetics in C.S.H. Laboratory*, C.S.H. New York, 119.

Sills & Stewart (1982), *J. Inst. Brew.* 88, 313-316.

15 Sills et al. (1984 a), *J. Inst. Brew.* 90, 311-314.

Sills et al. (1984 b), *APPL. Microb. Biotechnol.* 20, 124-128.

Southern (1975), *J. Mol. Biol.* 98, 503-517.

Stinchomb et al. (1980), *Proc. Natl. Acad. Sci. USA* 77, 4 559-4 563.

Struhl et al. (1979), *Proc. Natl. Acad. Sci. USA* 76, 1035-1039.

20 Taniguchi et al. (1983), *Nature* 302, 305-309.

Teather & Wood (1982), *Appl. Env. Microbiol.* 43, 777-780.

Towbin et al. (1979), *Proc. Natl. Acad. Sci. USA* 76, 4350-4354.

von Heijne (1983), *Eur. J. Biochem.* 133, 17-21.

Wilson & Ingledew (1982), *Appl. Environ. Microbiol.* 44, 301-307.

25 Zalkin & Yanowsky (1982), *J. Biol. Chem.* 257, 1491-1500.

Claims

30 1. A yeast cell of the genus Schwanniomyces,
characterized in that
 said yeast cell contains at least one expression cassette comprising :
 a) a first DNA sequence serving as a regulon,
 b) optionally a second DNA sequence coding for a signal peptide
 35 c) a third DNA sequence coding for a foreign protein, and
 d) optionally a fourth DNA sequence serving as a terminator.

2. The yeast cell according to claim 1,
characterized in that
 40 said first and/or second and/or fourth DNA sequence are derived from a gene coding for an amylolytic enzyme.

3. The yeast cell according to claim 2,
characterized in that
 said amylolytic enzyme is an α -amylase or glucoamylase.

45 4. The yeast cell according to any of claims 1 to 3,
characterized in that
 said first and/or second and/or fourth DNA sequences are derived from a yeast of a member of the genera Debariomyces, Saccharomycopsis, Lipomyces, Pichia, or Saccharomyces, preferably Schwanniomyces.

50 5. The yeast cell according to claim 4,
characterized in that
 said yeast is Schwanniomyces occidentalis.

6. The yeast cell according to any of claims 1 to 5,
characterized in that
 said first DNA sequence is comprised by a 1.8 kb BglIII-XbaI-fragment preceding the structural gene coding for the Schwanniomyces α -amylase or the 1.3 kb BamHI-PvuII fragment preceding the structural gene coding for the Schwanniomyces occidentalis glucoamylase.

55 7. The yeast cell according to claim 5,
characterized in that

said first DNA sequence is comprised by part or all of the DNA sequence corresponding to bases -1 to -540 of the region preceding the structural gene coding for α -amylase or part or all of the DNA sequence corresponding to bases -1 to -320 of the region preceding the structural gene coding for glucoamylase.

8. The yeast cell according to any of claims 1 to 5,

5 characterized in that

said first DNA sequence corresponds to one of the following regulons :

a) ADH1,ADH2, PDC1, GAL1/10, PGK, GAPDH, the regulon being preferably obtained from Saccharomyces cerevisiae

b) LAC4 the regulon being preferably obtained from Kluyveromyces lactis, or

10 c) corresponding regulons, preferably obtained from Schwanniomyces.

9. The yeast cell according to any of claims 1 to 5,

characterized in that

said first and/or fourth DNA molecule corresponds to a viral promoter and/or terminator, preferably from E. coli phage T7.

15 10. The yeast cell according to any of claims 1 to 9,

characterized in that

said second DNA sequence is coding for all or part of at least one of the following peptides :

a) MetArgPheSerThrGluGlyPheThrSerLysValAlaAlaAlaLeuAlaPheSerArgLeuValSerAla;

b) MetArgPheSerThrGluGlyPheThrSerLysValAlaAlaAlaLeuAlaPheSer-

20 ArgLeuValSerAlaGinProllePheAspMetArg;

c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAla;

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAlaAlleGlnAlaAlaProAla;

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu

ValSerAlaAlleGlnAlaAlaProAlaSerSerGlySerSerAlaSerAla

25 11. A yeast cell according to claim 10,

characterized in that

the DNA sequence coding for said peptide corresponds to part or all of one of the following DNA sequences:

a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT TTAGCATTCTCAAGATT-
30 GGTATCTGCT,

b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT TAGCATTCTCAAGATT-
GGTATCTGCTCAACCGATTATTTTGTGAGATGAGA;

c) ATGATTTCTGAAGCTGATTAAAAGTATAGTAAATTGGTTGGGATTA GTTAGTGCT;

d) ATGATTTCTGAAGCTGATTAAAAGTATAGTAAATTGGTTGGGATTA GTTAGTGCTATCCAAG-
35 CAGCCCCCTGCC;

e) ATGATTTCTGAAGCTGATTAAAAGTATAGTAAATTGGTTGGGATTA GTTAGTGCTATCCAAG-
CAGCCCCCTGCCCTTCGATTGGATCTAGTGCTTC GCA.

12. The yeast cell according to any of claims 1 to 11,

characterized in that

40 said third DNA sequence coding for a foreign protein is a natural or synthetical DNA sequence encoding a cellulase, interleukine, insulin like growth factor, interferon, lymphokine, human growth factor, nerve growth factor, aprotinin, insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines or human granulocyte/macrophage colony stimulating factor.

13. The yeast cell according to any of claims 1 to 12,

45 characterized in that

said fourth DNA sequence originates from a yeast of a member of the genera Saccharomyces, Pichia, Hansenula, preferably Schwanniomyces.

14. The yeast cell according to claim 13,

characterized in that

50 said fourth DNA sequence corresponds to part or all of the terminator of the Schwanniomyces α -amylase gene or part or all of the terminator of the Schwanniomyces glucoamylase gene.

15. The yeast cell according to claim 14,

characterized in that

said terminator is comprised by nucleotides 1537 to 1740 of the Schwanniomyces α -amylase gene or
55 nucleotides 2875 to 3320 of the Schwanniomyces glucoamylase gene.

16. The yeast cell according to any of claims 1 to 15,

characterized in that

said first and/or second and/or third and/or fourth DNA sequence has been modified by insertion or deletion

or substitution of one or more nucleotides while the biological activity of the DNA sequence is retained or improved.

17. The yeast cell according to any of claims 1 to 16,
characterized in that
 - 5 said expression cassette is carried by a circular or linear vector, which optionally comprises one or more selective marker genes.
 18. The yeast cell according to claim 17,
characterized in that
 - 10 said vector comprises any of the following selective marker genes : TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, preferably obtained from Saccharomyces, Hansenula Pichia or Schwanniomyces, AMY1 or GAM1, preferably obtained from Schwanniomyces.
 19. The yeast cell according to any of claims 17 and 18,
characterized in that
 - 15 said vector further comprises a DNA sequence capable of controlling autonomous replication and stable maintenance of the vector in the host organism.
 20. The yeast cell according to claim 19
characterized in that
 - 20 said DNA sequence is derived from the genome of a member of the genera Saccharomyces, Pichia, Hansenula or Kluyveromyces, preferably Schwanniomyces.
 21. A yeast cell according to any of claims 19 or 20,
characterized in that
 - 25 said DNA sequence is selected from the autonomously replicating sequences ARS, preferably an ARS derived from the respective host organism.
 22. The yeast cell according to claim 21,
characterized in that
 - 30 said DNA sequence is SwARS1 or SwARS2, capable of controlling autonomous replication and stable maintenance of the vector in Schwanniomyces.
 23. The yeast cell according to claim 22,
characterized in that
 - 35 said DNA sequences are comprised on a Clal-BamHI fragment (SwARS1) or a EcoRI-BamHI-fragment (SwARS2) of Schwanniomyces genomic DNA.
 24. The yeast cell according to any of claims 19 to 23,
characterized in that
 - 40 said DNA sequence has been modified by deletion, insertion or substitution of one or more nucleotides, while retaining or improving its biological function.
 25. The yeast cell according to any of claims 1 to 24,
characterized in that
 - 45 said yeast cell comprises any DNA-sequence homologous to genomic Schwanniomyces DNA, preferably homologous DNA-sequences flanking the expression cassette.
 26. The yeast cell according to any of claims 1 to 25,
characterized in that
 - 50 said expression cassette is integrated into the genome, optionally in multiple copies, preferably via homologous recombination.
 27. The yeast cell according to any of claims 1 to 26,
characterized in that
 - 55 said yeast cell is derived from the genus Schwanniomyces and is capable of glycosylation.
 28. A process for the production of a polypeptide, wherein a yeast host organism is cultivated under suitable conditions and the polypeptide is recovered in a manner known per se,
characterized in that
 - 50 said yeast host organism is a yeast cell according to any of claims 1 to 27.
 29. The process according to claim 28
characterized in that
 - 60 yeast cells of the species Schwanniomyces occidentalis are used as a host organism and are preferably cultivated on starch,dextrin, maltose and/or plant biomass.
 30. The process according to any of claims 28 or 29,
characterized in that
 - 65 yeast cells of the genus Schwanniomyces are used for the production of single cell protein.
 31. The process according to any of claims 28 to 30,

characterized in that

said polypeptide is α -amylase or glucoamylase.

32. The process according to claim 28,

characterized in that

5 the signal sequence of a yeast α -amylase or glucoamylase gene is used for the production and secretion of foreign protein in yeast of the genera Saccharomyces, Kluyveromyces, Hansenula, Pichia or Schizosaccharomyces.

33. Use of a yeast cell according to any of claims 1 to 27 for transformation and expression of foreign genes.

10

15

20

25

30

35

40

45

50

55

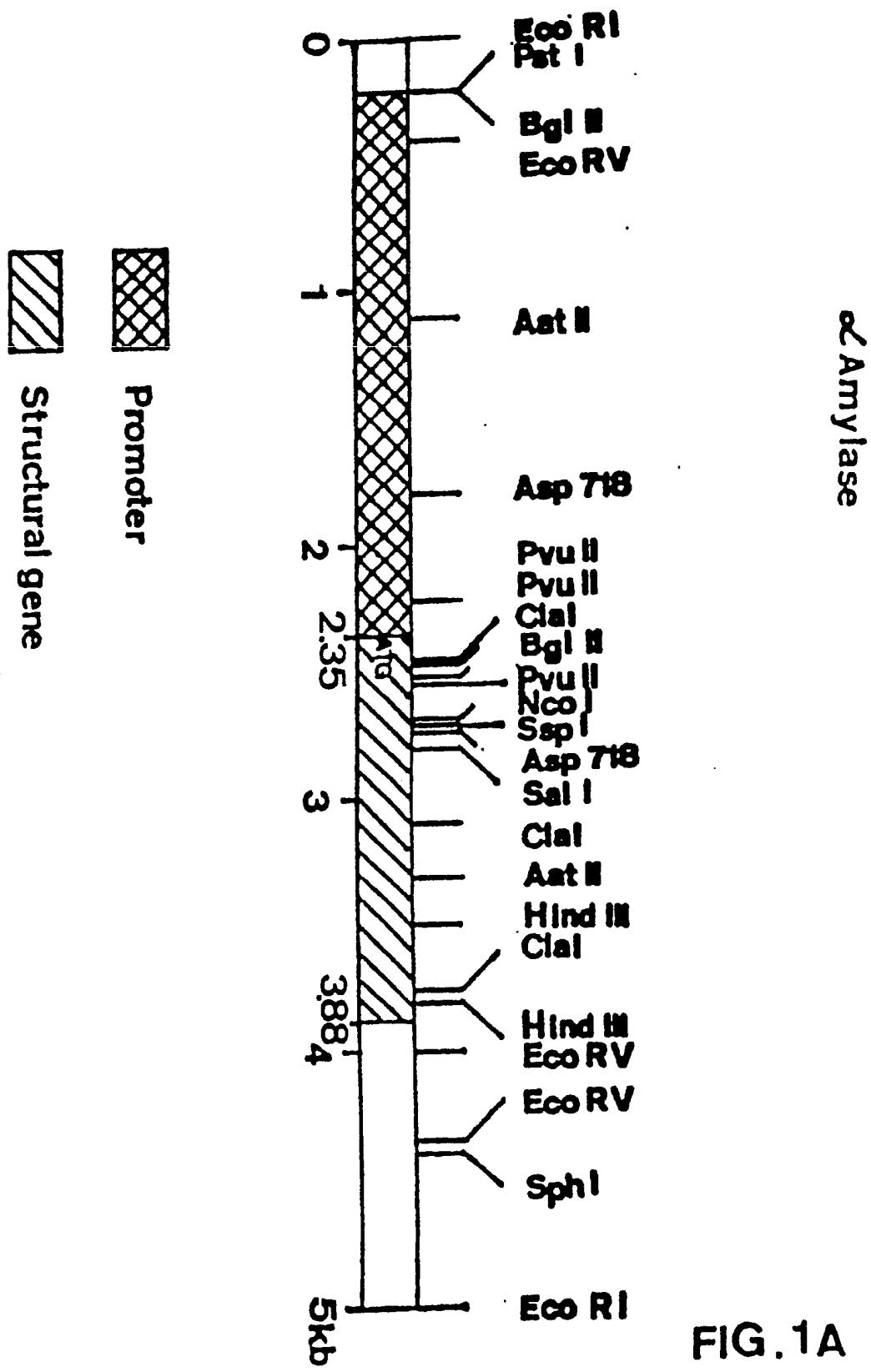


FIG. 1A

FIG 1B,1

GGTACCTGA

-540 GCTAAATTAA GAAACCGGCTA TAGATCTGGT TGTCTAAAGA AGAGATAATG AAGAAACAA
 -480 TTAACCCGAGC ACTCTTATTA AGTTTTTTTC TATTTTCTTT TGCTCTACT TCAATAATT
 -420 ATCTAAATTG TATTGGCGT TAGATCAGAA TGTACTGATA ACAGAGAGTA TTATCATACA
 -360 CTTATGGATT TCAAAAGGGCG GAATCAAAAG CATACTAAT CAAACCCCTTG GTTATTTGAT
 -300 GCAATTAAAGG TTGTAGTCGT TCTTACCGAT CCATCATTAT ACCCCACACG GTTTCATGGT
 -240 ATGTAGGGTGT TTCAATAGTG AAGTACAATG AATGTGTTTG TAATGCTGTA TGTGGATCAG
 -180 TAATTATGTT AACAAATTAA GTCTGAAAT TTATTAAAT TTTACCTACA AATTAAGGCC
 -120 AAATCCAATC GAAGGTGCCG CCCAGCTGGT GTATAAATTAA CTCTGAAAT TCAAGTTGAA
 -60 CGTTGATCTC TCTAAAGCAA AGCTGTTATT CTACAATACT AAATAAAATA AAAGCAAGAC

310	320	330	340	350	360
GTTGAGCAAAATTCCCTGACAATACTGCTTATGGTTACCATGGTTATTGGATG					
ValValGluGlnIleProAsnThrAlaTyrGlyTyrAlaTyrHisGlyTyrTrpMet					
370	380	390	400	410	420
AAAATATTGATGAATTGAAACACTAAATTGGTACCGCTGATGAATTGAAACAAATTAGCT					
LysAsnIleAspGluLeuAsnThrAsnPheGlyThrAlaAspGluLeuLysGlnLeuAla					
430	440	450	460	470	480
AGCGAATTGCATTCCAGAAGCATGTATTGATGGTCGACCGGTTACAAACCATTATGCT					
SerGluLeuHisSerArgSerMetLeuLeuMetValAspValValTyrAsnHisTyrAla					
490	500	510	520	530	540
TGGAACGGAGATGGTTCAAGCGTAGATTATTCTAGTTTCACTCCATTCAACAAATCT					
TrpAsnGlyAspGlySerSerValAspTyrSerPheThrProPheAsnGlnGinSer					
550	560	570	580	590	600
TACTTCCACGATTATTGTTGATTACAATTATAATGATCAAACCAATGTTGAAGATTGCT					
TyrPheHisAspTyrCysLeuIleThrAsnTyrAsnValGluAspGlySerGlySer					
610	620	630	640	650	660
TGGGAAGGTGATACTGAAAGTCTCCAGATTCCAGGATAATTGAAAGTTATA					
TrpGluGlyAspThrGluValSerLeuProAspLeuSerThrGluAspAsnGluValIle					

670 680 690 700 710 720
 GGAGTATTCAAACCTGGGTCAAGATTGTTCAAAACTATTCAATCGATGGTTAAGA
 GlyValPheGlnThrTrpValSerAspPheValGlnAsnTyrSerIleAspGlyLeuArg
9
 730 740 750 760 770 780
 ATTGATAGTGCAAGGCACGGTAGATACCGCTTCATTAAACGAAAGTTGAGGACGCTCTGGT
 IleAspSerAlaLysIleValAspThrAlaSerLeuThrLysPheGluAspAlaSerGly
 790 800 810 820 830 840
 GTTATAACTTAGGTGAAGTTTATCAAGGAGATCCAACCTATACTTGTCCATATCAGAAAT
 ValTyrAsnLeuGlyGluValTyrGlnGlyAspProThrTyrThrCysProTyrGlnAsn
 850 860 870 880 890 900
 TATATGAAAGGAGTTACCAACTATTCAACTTACATTACTATCCAGTATATAAGATCTTCAGTGTAT
 TyrMetLysGlyValThrAsnTyrProLeuTyrTyrProValTyrArgPhePheSerAsp
 910 920 930 940 950 960
 ACTTCGGCGACTTCCAGTTAACCTCAATTGATCTCCACGGTTACAGTCATCTTGTTCG
 ThrSerAlaThrSerSerGluLeuThrSerMetIleSerThrLeuGlnSerSerCysSer
11
 970 980 990 1000 1010 1020
 GACGTCTCTTGTGGAAACTTTATTGAAACCATGACCAAGTTAGATTCCATCAGTT
 AspValSerLeuLeuGlyAsnPheIleGluAsnHisAspGlnValArgPheProSerVal
 1030 1040 1050 1060 1070 1080
 ACCCTCAGACACATCCCTGATTAAAGAAATGACATGGCTTTATAAATTTGGGTGATGGTATC
 ThrSerAspThrSerLeuIleLysAsnAspMetAlaPheIleIleLeuGlyAspGlyIle

1090 1100 1110 1120 1130 1140
 CCAATTATTATGCCAAGAACATGGTGGTCAATGGTGGTCCGATCCTGCCAATAGA
 ProIleLeuTyrTyrGlyGlnGluGlnGlyLeuAsnGlySerAspProAlaAsnArg
 1150 1160 1170 1180 1190 1200
 GAAGCTTTATGGTTAAGGTGATATAACCGATTCAAGAAATACTACCGAGCTAATCAGTAAGA
 GluAlaLeuTyrTyrGlySerGlyTyrAsnThrAspSerGluTyrTyrGluLeuIleSerLys
 1210 1220 1230 1240 1250 1260
 CTAATCAAATAAGGAAATCAAGCTTAAAGAAGGGATTCTGCCTATTCAACTTACAAATCC
 LeuAsnGlnIleArgAsnGlnAlaIleLysAspSerAlaTyrSerThrTyrLysSer
 1270 1280 1290 1300 1310 1320
 TCAGTTGTTCTCTTCAAGACCCTTAACTAGCCACTAGGAAGGGTAGCGATGCTAATCAA
 SerValAlaSerSerAspHistYrIleAlaThrArgLysGlySerAspAlaAsnGln
 1330 1340 1350 1360 1370 1380
 TGATTTCATTTTAAATTAGGTTCAAAACGGGTCAACAGGATATTACTGTCAAGCAAC
 LeuIleSerIlePheAsnAsnLeuGlySerAsnGlySerGlnAspIleThrValSerAsn
 1390 1400 1410 1420 1430 1440
 ACCGGCTATTCTAGTGGTGTATAAAAGTTATCGATATTTCCTGCCAATTCCGTTTAGCT
 ThrGlyTyrSerSerGlyAspLysValIleAspIleSerCysAsnSerValLeuAla
 1450 1460 1470 1480 1490 1500
 GGTGACTCCGGAAAGCTTATCTGTATCAAATTCTGGTGGAAATGCCACAAAGTTACGGCTCG
 GlyAspSerGlySerValSerValSerValSerGlyGlyMetProGlnValTyrAlaPro

1510 1520 1530 1540 1550 1560
TCC TCT GTT CTT CGGGATCTGGCATCTGCAATAGATTGATCCAGCGCTAACCCCT
Ser Ser Val Leu Ser Gly Ile Cys Asn Glu
1570 1580 1590 1600 1610 1620
TTTTAGGAAACGACAAGTTTATTTAGAAAAAGTTTCTAAGAATGGTCAAAACAAAGTTCT
1630 1640 1650 1660 1670 1680
TTATTACTCTATGTCCTGGATATCTGTTCAATGTTCTGACTCCACATTCCCTCAT

1690 1700 1710 1720 1730 1740
GTTAGTTCTTCTTGTGTCGATCTCTAAGTTTATTCTTAATTAAATCCAAA

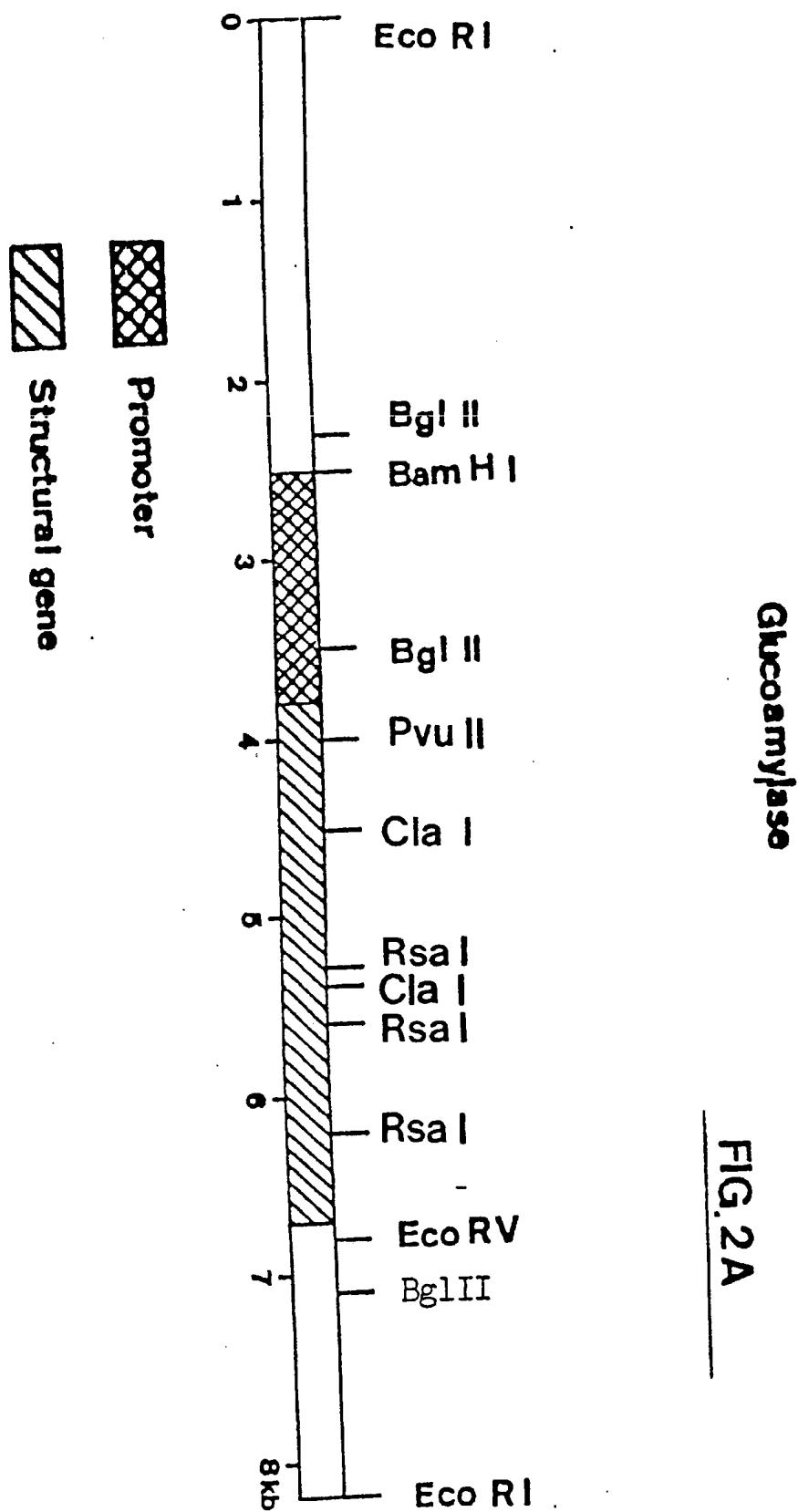


FIG. 2B₁

-310
AGATCTACATTAAACCCAG

-290
TCTACTCCAGATATTGGAGTATAACCCATTCTACCCATTAACTCTAACTGCA

-280
AATTTCAAAAGGATTTCGAGGAAATTCTTCTAAATAACGAAAGTGTATTGGTATTCA

-150
ATTACTACGGAAACTACTCATATGGTAGTAACTGTTGGTAAATGCGCAATTGTAATT

-100
GCCGAAGTTATAATAGTTGGCAAAACTGGAGAATTTCATTATTGGGAAATAATAAA

-50
TAAGGCCAAGTATCCATTGAAATTAAATGAACTCATGACTGTATTATAACCAAGCAAG

-1

10 20 30 40 50 60
 ATGATTCTCTGAAAGCTGATTAAAGTATAGTAAATTGGTTGGGATTAGTTAGGCTATC
 Met Ile Phe Leu Ile Lys Ser Ile Val Ile Gly Leu Glu Val Ser Ile Ile
 ★
 70 80 90 100 110 120
 CAAGCAGCCCCCTGCCCTTCGATTGGATCTAGTGCTTCAGCATCTAGTTCAAGTGAGAGT
 Gln Ile Ala Pro Ala Ser Ser Ile Gly Ser Ser Ser Ser Glu Ser
 ★
 130 140 150 160 170 180
 TCTCAGGGCTACAATTCCCAATGATGTAACATTAGTGTAAACAAATTCCCTAAATATCTT
 Ser Gln Ile Pro Asn Asp Val Ile Thr Leu Glu Val Ile Gln Ile Pro Asn Ile Phe
 Ser Gln Ile Thr Ile Pro Asn Asp Val Ile Asn Ile Asn Ile Asn Ile Asn
 190 200 210 220 230 240
 AATGACTCTGCTGATGCTTAATGCAGCTAAAGGGTATGACTTGCTAAATGTTACT
 Asn Asp Ser Ile Val Ile Asp Val Ile Asn Ile Asn Ile Asn Ile Asn Ile Asn
 250 260 270 280 290 300
 AATACTCCAAAGAGGATTAAACTTAAACCGGGTATCTTAAAGAAGCTACCAATATTATGGT
 Asn Thr Pro Arg Gln Ile Leu Thr Gly Ile Leu Ile Val Ile Asn Ile Asn Ile Asn Ile Asn
 310 320 330 340 350 360
 TATGATTGATTAAACTTAAACTGTTGAATAACCAAGCTGATGCTGATTAAACGTT
 Tyr Asp Phe Asp Tyr Leu Asn Leu Thr Val Glu Tyr Gln Ile Asp Thr Arg Glu Asn Val
 370 380 390 400 410 420
 CATATTGAAACCAACTGATTCTGATGTATTGTTTACCAAGAGCATTAGTTGTTAAAG
 His Ile Glu Pro Thr Asp Val Phe Val Leu Pro Glu His Leu Val Ile Val Ile

430 440 450 460 470 480
 CCACCTGGAGGTGATGCCACAATTAACTTGGACAAATTCCGATTGGTTTCGAA
 ProLeuValGluGlyAspAlaGlnSerTyrAsnPhoAspAsnSerAspLeuValPheGlu

 490 500 510 520 530 540
 TACTCTAACTGACTTCTCCTTTGAAGTTATTAGATCATCTACTAAAGTTTTATT
 TyrSerAsnThrAspPheSerPheGluValIleArgSerSerThrLysGluValLeuPhe

 550 560 570 580 590 600
 TCTACTAAAGGTAAATCCATTGGTTTCAAAATCAATTCAATTCAATTGTCATTG
 SerThrLysGlyAsnProLeuValPheSerAsnGlnPheIleGlnPheAsnSerSerLeu

 610 620 630 640 650 660
 CCAAGAACCATGGTTATTACTGGCTTGGTGAATCTTACCGGTTAGTTAACGAAACCA
 ProLysAsnHisValIleThrGlyLeuGlyGluSerIleHisGlyLeuValAsnGluPro

 670 680 690 700 710 720
 GGTAGCGTTAACATTATTGCTTAATGATGTTGGTGAATCCAAATCGATGGTAATTTAT
 GlySerValLysThrLeuPheAlaAsnAspValGlyAspProIleAspGlyAsnIleTer

 730 740 750 760 770 780
 GGTGTCATCCAGTTTATCTTGATCAAAGATATGACACTGAAACTACCCATGCTGTTTAT
 GlyValHisProValIleTerLeuAspGlnArgTyrAspThrGluThrThrHisAlaValIleTer

 790 800 810 820 830 840
 TGGAGAACTTCTGCTTATTCAAGAAGTATTAAATCGGTGAGGAATCTTACTTGGAGAGCT
 TrpArgThrSerAlaIleGlnGluValLeuValGlyGluGluSerIleThrTrpArgAla

850 860 870 880 890 900
 CTTTCAGGTGTTATTGATTTACCTTCTTAGTGGTCTTACACCAAAAGATGCCATTCAA
 LeuSerGlyValLeuAspLeuTyrPhePheSerGlyProThrProLysAspAlaIleGln
 910 920 930 940 950 960
 CAGTATGTCAAAGAGATTGGTTACCAAGCTTTCCAAACCTACTGGCTTAGGTTACCAT
 GluTyrValLeuGluLeuGlyLeuAspAlaPheGlnProTyrTrpSerLeuGlyTyrHis
 970 980 990 1000 1010 1020
 CAATGTAGATGGGGTTACGATACTATCGAAAATTATCTGAAGTTGTTGAAACTTCAG
 GluCysArgTrpGlyTyrAspThrIleGluLysLeuSerGluValLeuGluAsnPhelys
 1030 1040 1050 1060 1070 1080
 AAATTAAATTCCATTAGAAAACCTATCTGGTCAGACATGGATTACATGGACTCTTAA
 LysPheAsnIleProLeuGluThrIleTrpSerAspPheAspIleAspTyrMetAspSerTyrLys
 1090 1100 1110 1120 1130 1140
 GATTTCACCTTATGATCCACAGATTCCCACTAGATGAATATCGTAAATTCCCTTGATGAG
 AspPheThrTyrAspProHisArgPheProLysAspGluLeuAspGlu
 1150 1160 1170 1180 1190 1200
 TTGACACAAATAATCAACACTATGTTCTATTGGATGCTGCTATTACGTTCCAAAC
 LeuHisLysAsnAsnGlnHisTyrValProIleLeuAspAlaIleTyrValProAsn
 1210 1220 1230 1240 1250 1260
 CCAAAACAAATGCTAACGAAATACCAACCTTCCACTATGGTAATGAAACCGATGTC
 ProAsnAsnIleThrAspAsnGluIleTyrGlnProPheHisTyrGlyAsnGluThrAspVal

1270 1280 1290 1300 1310 1320
 TTCTTAAGAATCCAGATGGTTCAATTATGTTGGCTGTGGCAGGGTTACACTGTT
 PheLeuLysAsnProAspGlySerLeuTyr11>Gly11ValTyrGlnValThrLeuPhe
 1330 1340 1350 1360 1370 1380
 TCCAGATTCTTAGCAGAAACATTCAAGATATGGATTAAGTCATTAAAGATTGGTATGAA
 SerArgPheLeuSerArgLysHisSerAspMetAspLysVal11>LysAspTyrTyrGlu
 1390 1400 1410 1420 1430 1440
 TTAACCTCCCTTGTGGTATTGGCTGATATGAAATGAAATCTCATCTCTGTGTGGT
 LeuThrProPheAspGly11>TyrAlaAspMetAspGluValSerSerPheCysValGly
 1450 1460 1470 1480 1490 1500
 TCTTGGGTACTGGTAAATACCTTCGAAACCCAGGCAATCCCTCCATTACTGTTGGAAAGT
 SerCysGlyThrGlyLysTyrPheGluAlaAsnProAlaTyrProProPheThrValGlySer
 1510 1520 1530 1540 1550 1560
 AAAGCTACCTCTTATCCAGTTGGTTCGATGTTCTAACGGCATCTGAATGGAAATCTATT
 LysAlaThrSerTyrProValGlyPheAspValSerAsnAlaSerGluTrpLysSerIle
 1570 1580 1590 1600 1610 1620
 CAAGGCTCAATTCTGCTACTGCTAAGACTTCTTCAACTTCTGCTGCTCTCA
 GlnSerSerIleSerAlaThrAlaLysThrSerSerSerValSerSerSerSer
 1630 1640 1650 1660 1670 1680
 TCCACAAATCGATTATGAAACACTTCTGCTTCAAGGTAAAGGTAAATTATCCACCA
 SerThrIleAspTyrMetAsnThrLeuAlaProGlyLysGlyAsnIleAsnTyrProPro

1690 1700 1710 1720 1730 1740
 TATGCTATTACAAACATGCAAGGTGACTCCGATCTTGCTACTCTGAGTATCTCCAAAT
 TyrAlaIleTyrAsnMetGlnGlyAspSerAspLeuAlaThrHisAlaValSerProAsn

 1750 1760 1770 1780 1790 1800
 GCTCACACATGCTGATGGTACAGTTGAAATATTGATAATTCAAACTCTTATGGTTACTTGCAA
 AlaThrHisAlaAspGlyThrValGluTyrAspIleHisAsnLeuTyrGlyTyrLeuGln

 1810 1820 1830 1840 1850 1860
 GAAAATGCTACTTATCATGCATTATTGGAAAGTTTTCTAACAAAGAACCCATTCAATTGATT
 GluAsnAlaThrTyrHisAlaLeuLeuGluValPheProAsnLysArgProPhoMetIle

 1870 1880 1890 1900 1910 1920
 TCCAGATCAACCTTCCACGGCCTGGTTAAATGGACCGCATTGGGGTGGTGACAACACT
 SerArgSerThrPheProArgAlaGlyLysTrpThrGlyHisTrpGlyGlyAspAsnThr

 1930 1940 1950 1960 1970 1980
 GCTGATTGGCTTATGCTTACCTCTCTATCCCTCAAGCATTCTCAATGGGTATTGCTGGC
 AlaAspTrpAlaIleTyrPhoSerIleProBlnAlaPhoSerMetGlyIleAlaGly

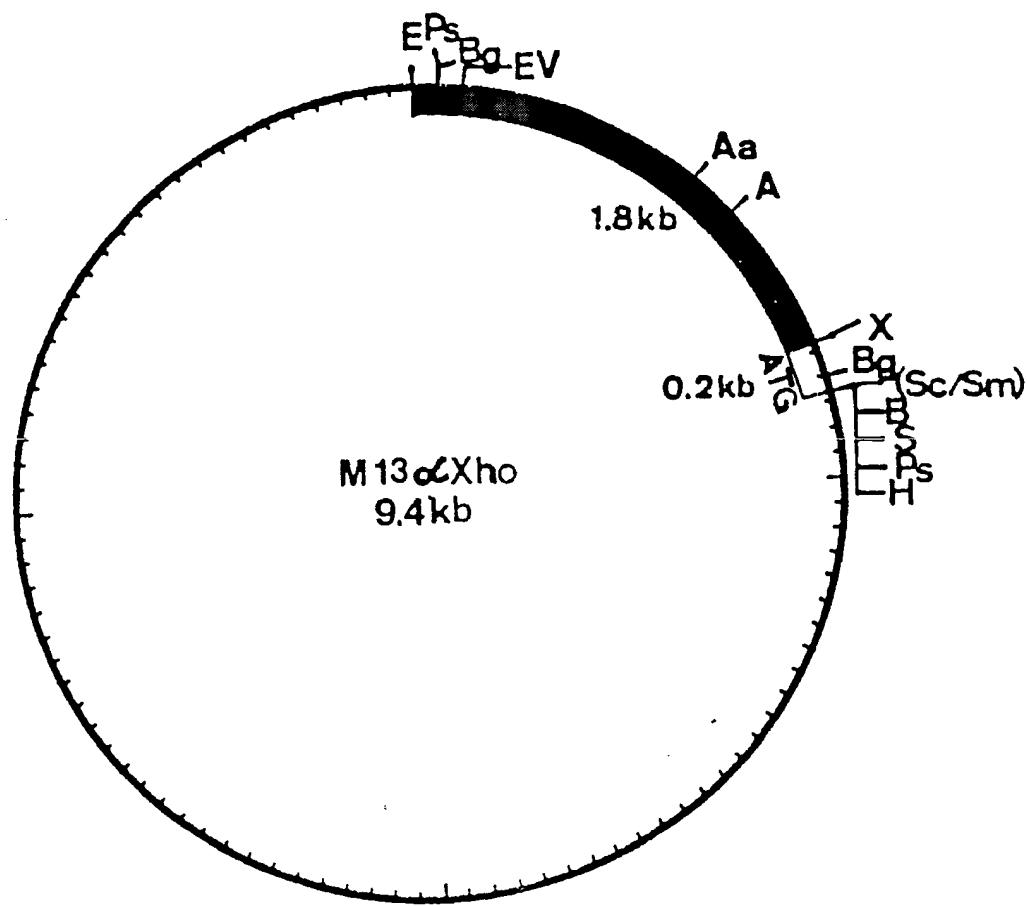
 1990 2000 2010 2020 2030 2040
 CTTCCATTGGTGGCGATGTTGGTTCAATGGTAATTCTGATTCTGAATTATGT
 LeuProPhoPheGlyAlaAspValIleGlyAspGlyAsnSerAspSerGluLeucys

 2050 2060 2070 2080 2090 2100
 TCAAGATGGATGCAATTAGGTTCTTCTTCCATTCTACAGAAACCAACTATTAGGT
 SerArgTrpMetGlnGlySerGlySerPheProPhoTyrArgAsnHisAsnTyrLeuGly

2b,6

2530 2540 2550 2560 2570 2580
 ACTGTTGCTGAGTCAGACAAATCCATTGGTTAATTGTCGCTTACGATGGC
 Thr Val Ile Glu Ser Arg Gln Asn Pro Phe Gly Leu Ile Val Ile Leu Asp Asn Asp Gly Y
 2590 2600 2610 2620 2630 2640
 AACGCTCAAGGTAGCTTACCTTGAATGGTGAATCATTAGTAGACTCTTCATTG
 Lys Ala Gln Gly Ser Ile Val Ile Leu Ile Val Ile Leu Asp Asn Asp Asn Asp Asn Asp Gly Y
 2650 2660 2670 2680 2690 2700
 TTGGTTAGTTCTCTGTTCTGATAACATTAACATCCATCTGGTGACTATAA
 Leu Val Ser Phe Ser Val Ser Asp Asn Thr Leu Ser Asp Asn Thr Leu Ser Asp Asn Thr Leu Ser Val
 2710 2720 2730 2740 2750 2760
 GCTGATCAACCTTGTGTTACCATCTAGGGGCTAACCAAAATCAGTT
 Ala Asp Gln Pro Leu Ile Asn Val Thr Ile Leu Ile Gln Val Gln Ile Val Gln Ile Val Gln Ile Val
 2770 2780 2790 2800 2810 2820
 AAATTGAAACGCTTAATGGTGAATGCTACCAAGAAATCAACCGTTTCGTTACTGGC
 Lys Phe Glu Asn Ala Asn Val Ile Leu Ile Val Ile Leu Ile Val Ile Leu Ile Val Ile Leu Ile Val Ile Val
 2830 2840 2850 2860 2870 2880
 TTAGATAAACCAACGGATGGTGCATTCTAAGGATTTCACCAATTACTGGTAATT
 Leu Asp Lys Tyr Thr Lys Asp Gln Ile Phe Ser Lys Asp Phe Thr Ile Thr Ile Thr Trp
 2890 2900 2910 2920 2930 2940
 TAACATCCACCTTAGTTCAATTCCATTCTTCTTCTGAAATTCTGAATTGAAA

2950 2960 2970 2980 2990 3000
 TTATTGAATGATATCATTAGTTCTTCAGCTTATGCTATGTTTCGATTAA
 ATGTTAAAGTTTATGTTTATGTTGTTGAGTTGATAAAATATAGCAA
 3010 3020 3030 3040 3050 3060
 TACATCGAAAATTGCGATGAAATTTCAGCTCATTAGAAATGTAGTCATTAGT
 3070 3080 3090 3100 3110 3120
 CACATGGACCACTATAACAAACAACAAACTATTCCAGAAAATATGTAAAGGATA
 3130 3140 3150 3160 3170 3180
 TAGATCATAAATTCTTATTGACTTTGTTTAAACATAAGTTACATAAGGAATATTG
 3190 3200 3210 3220 3230 3240
 TTTACTACAAACCATGGTCTGTAAAGAAGCAGACGAGGGCTATGTGTTGGTTGG
 3250 3260 3270 3280 3290 3300
 CCGCAATACTAGTTACAAAG



— M13

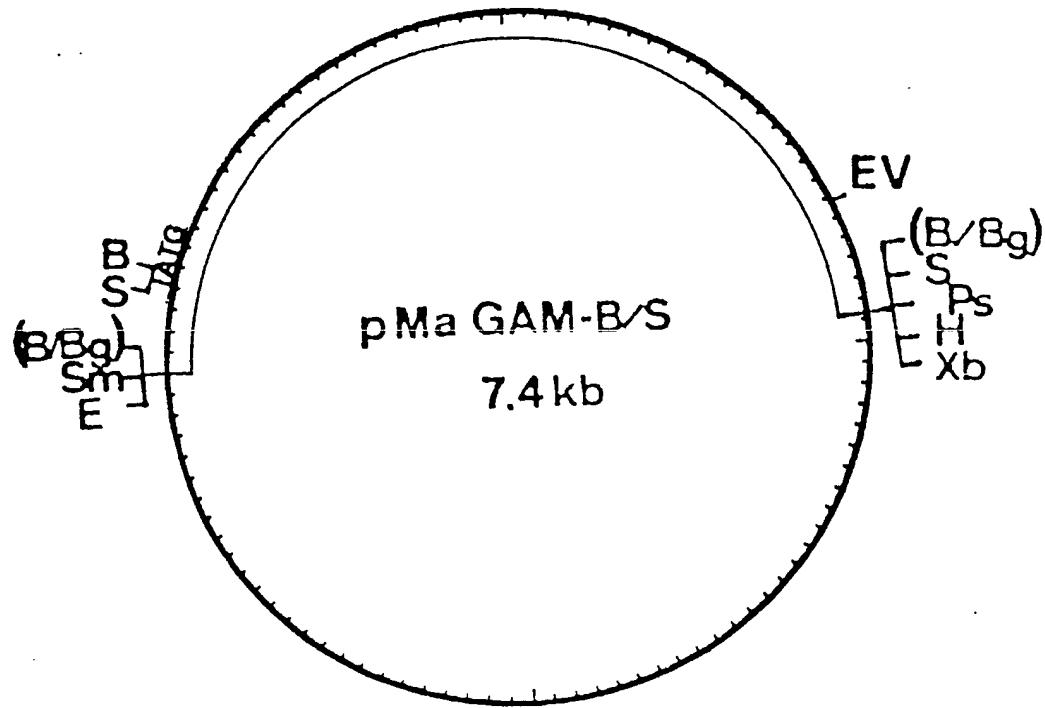
□ α Amylase
0.2kb of the structural gene

■ α Amylase-Promoter

E Eco RI
Ps Pst I
Bg Bgl II
EV Eco RV

Aa	Aat II
A	Asp 718
B	Bam H I
Sm	Sma I
S	Sal I
H	Hind III
Sc	Scal I
X	Xho I

FIG. 3



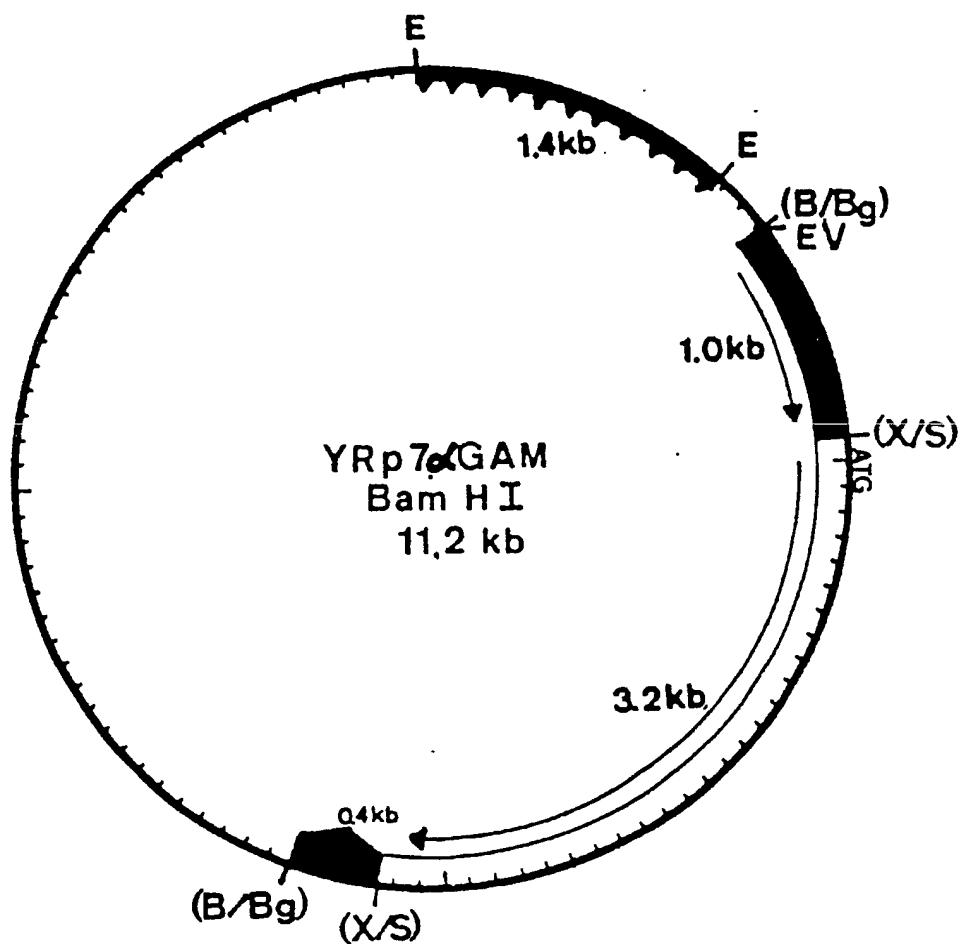
□ Glucoamylase

— pMac5-8
Kramer et al. 1984

EV Eco RV
B Bam H I
Bg Bgl II
S Sal I
Ps Pst I

H Hind III
Xb Xba I
E Eco RI
Sm Sma I

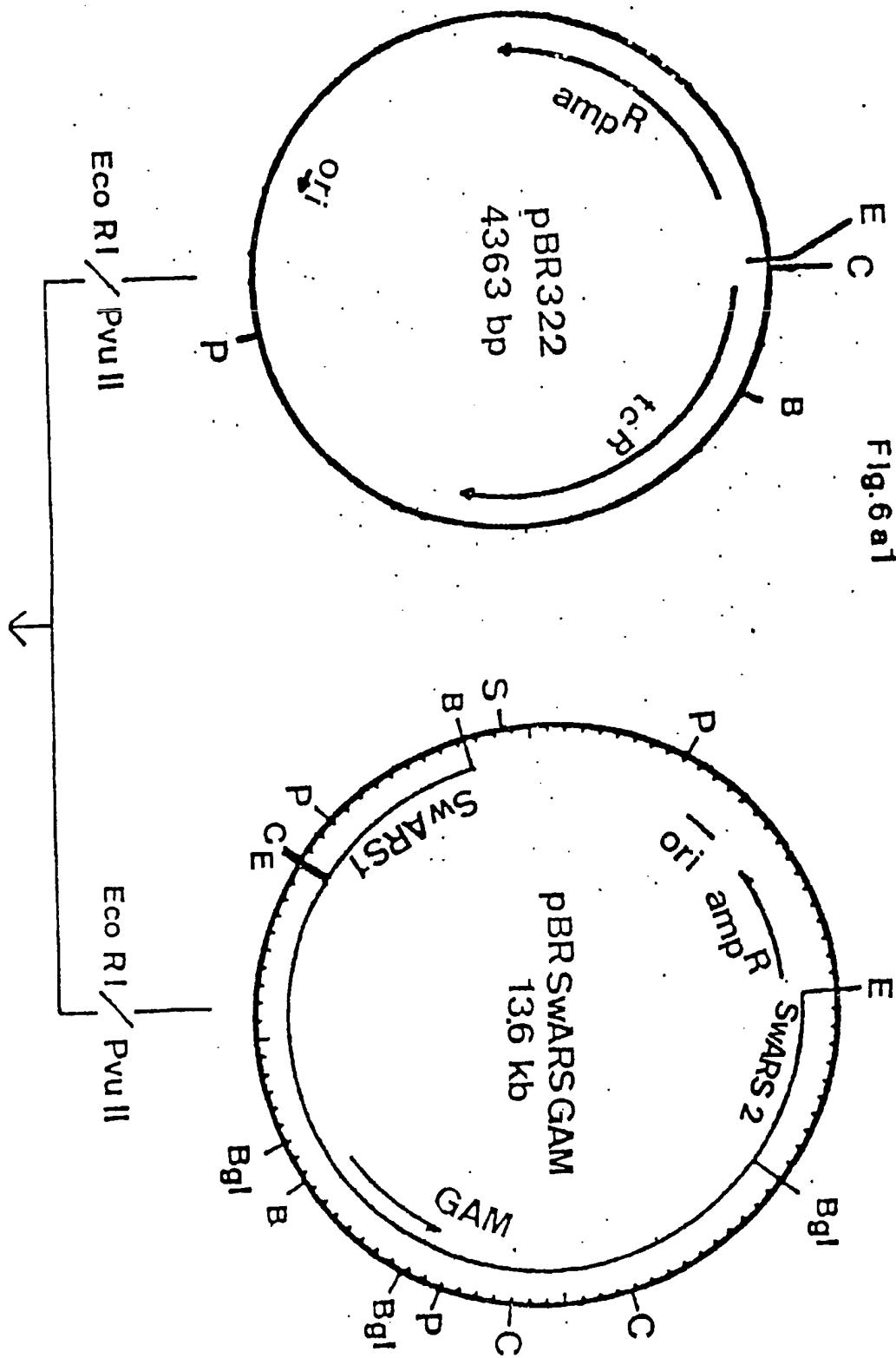
FIG.4



■■■■	TRP/ARS	E	Eco R I
■■■■	α Amylase-Promoter	B	Bam HI
□□□□	Glucoamylase structural gene	Bg	Bgl II
—	YRp7 Struhl et al. 1979	EV	Eco RV
■■■■	α Amylase 0.4 kb of the structural gene	X	Xho I
		S	Sal I

FIG. 5

Fig. 6 a1



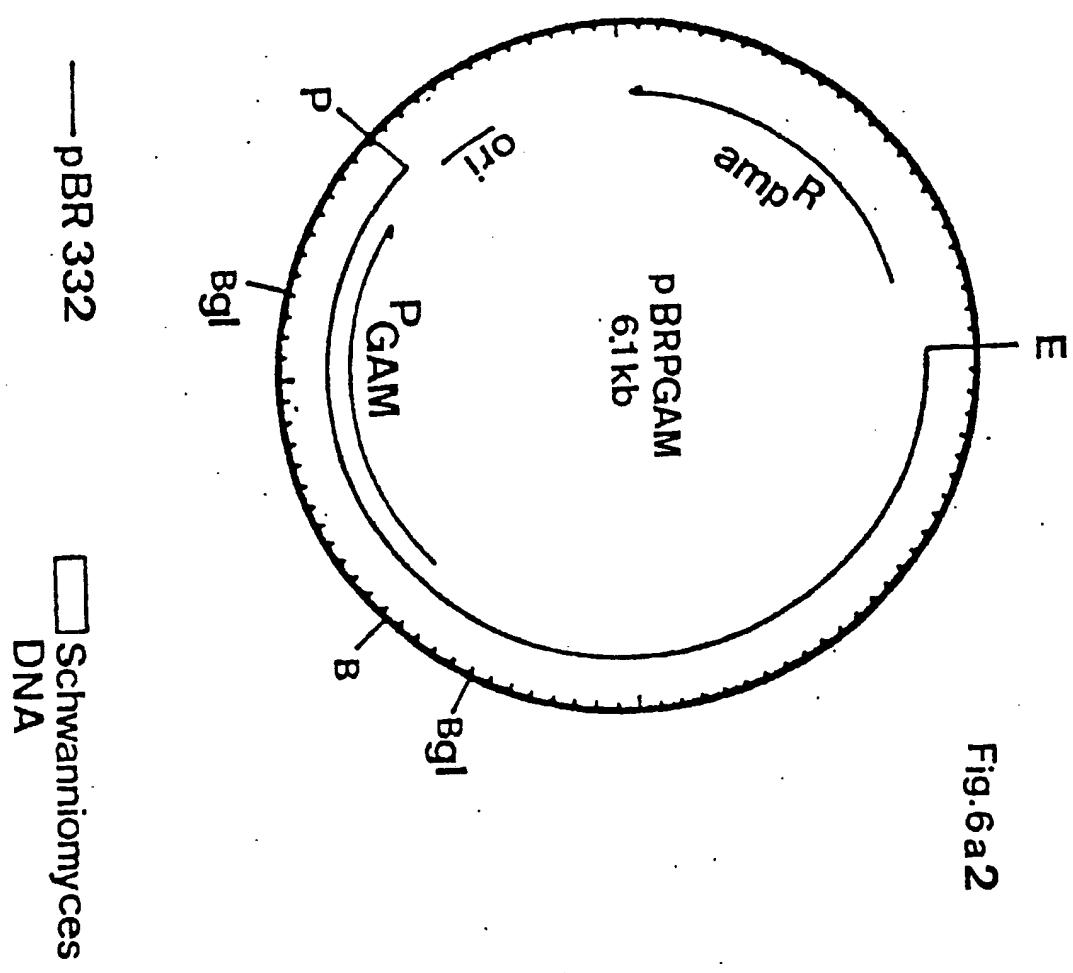


Fig.6a2

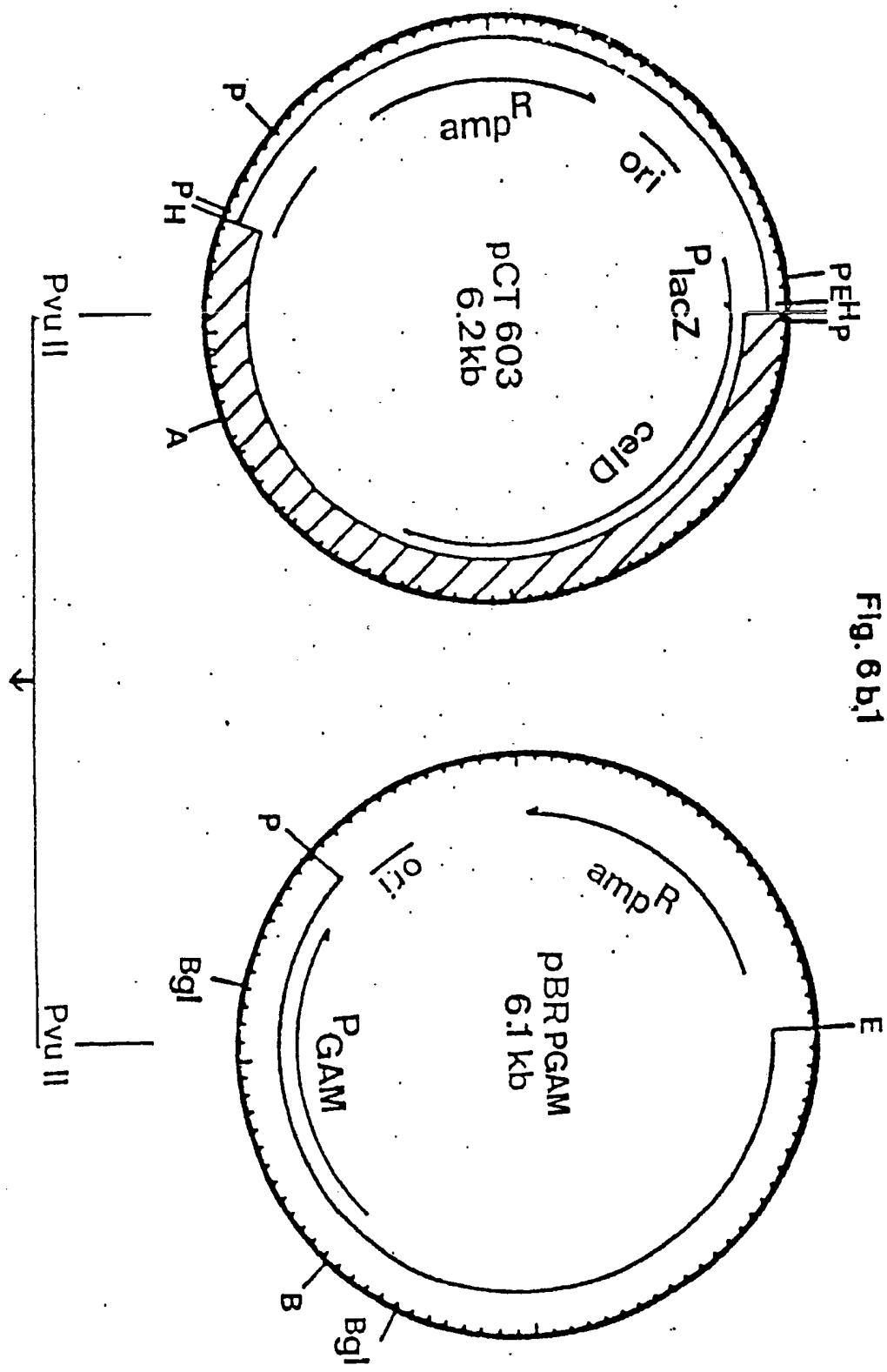


Fig.6b,2

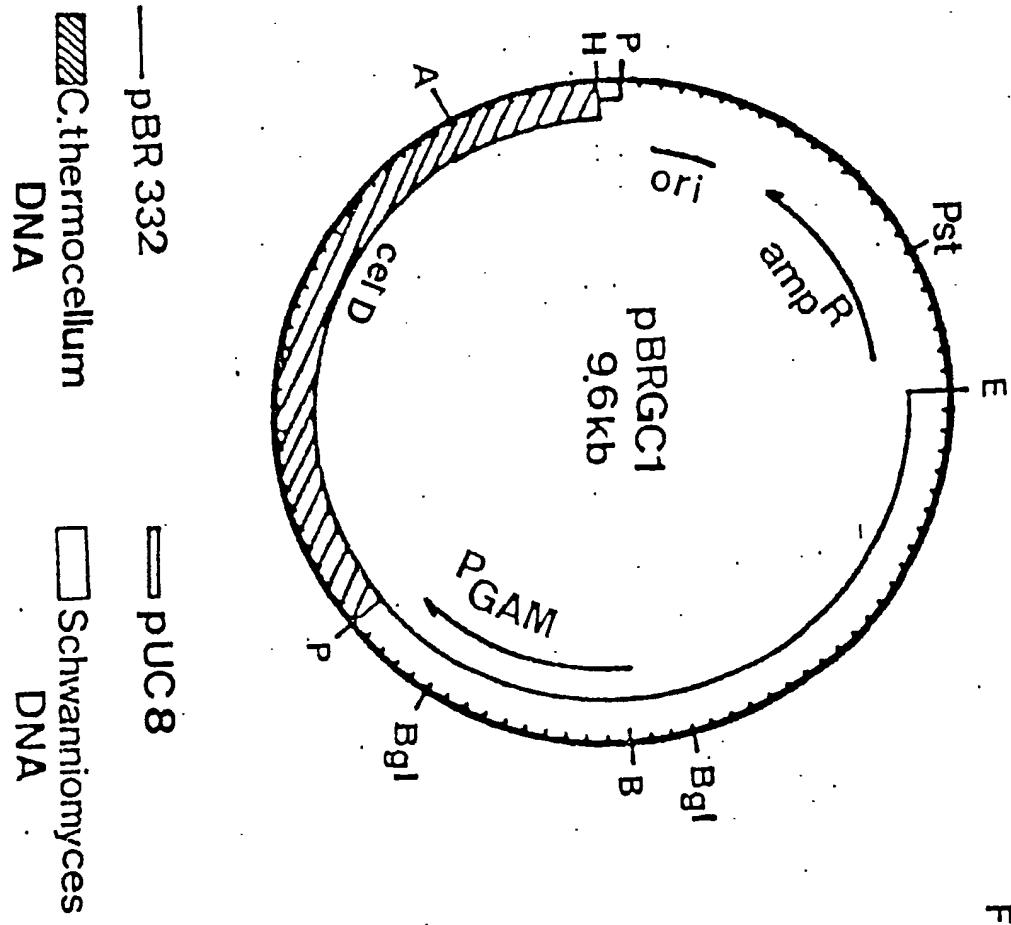
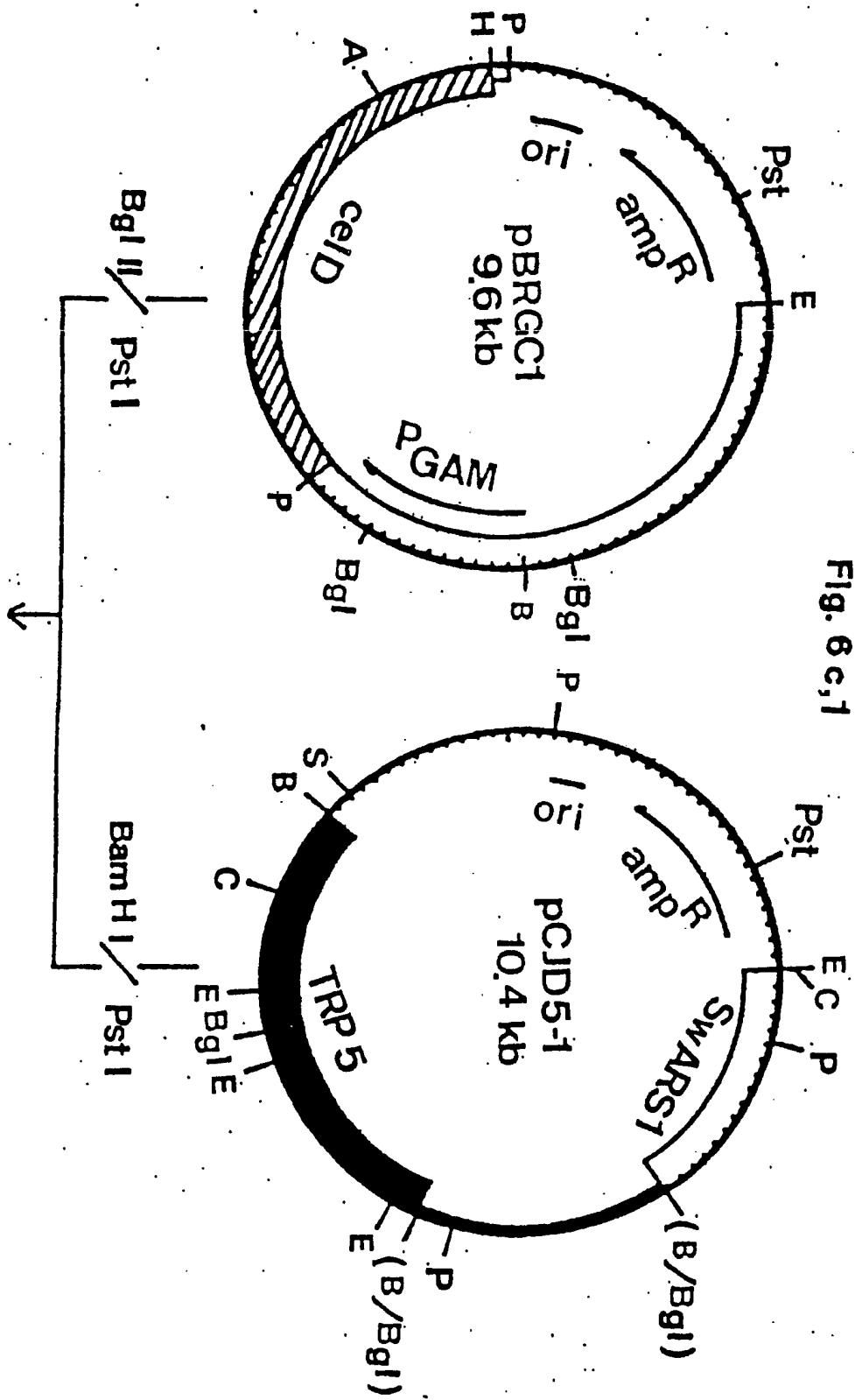


Fig. 6 c,1



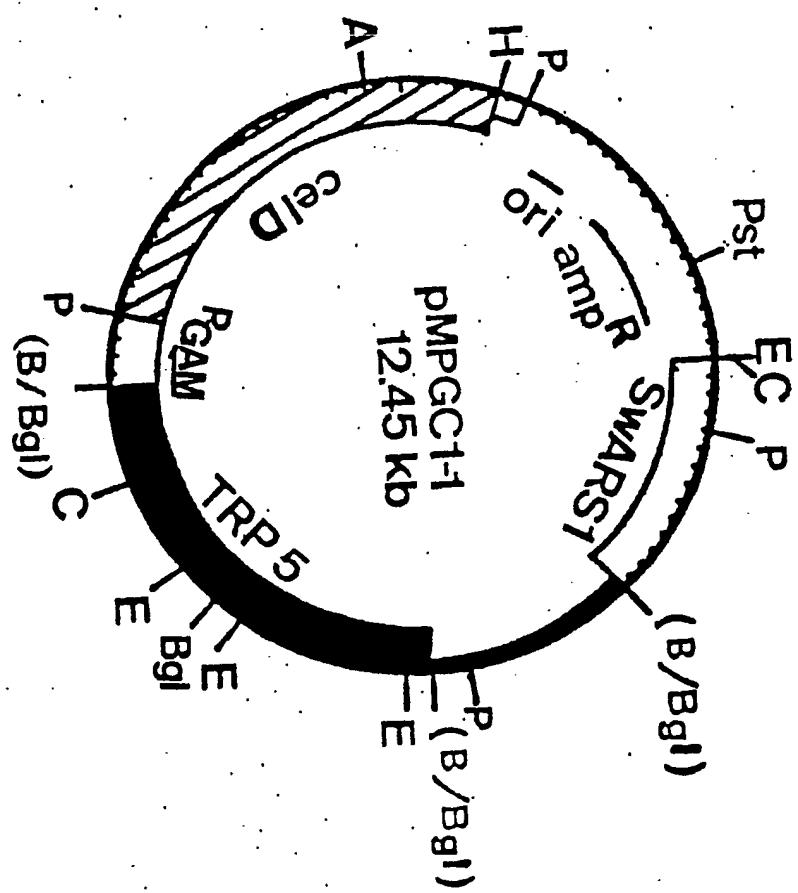
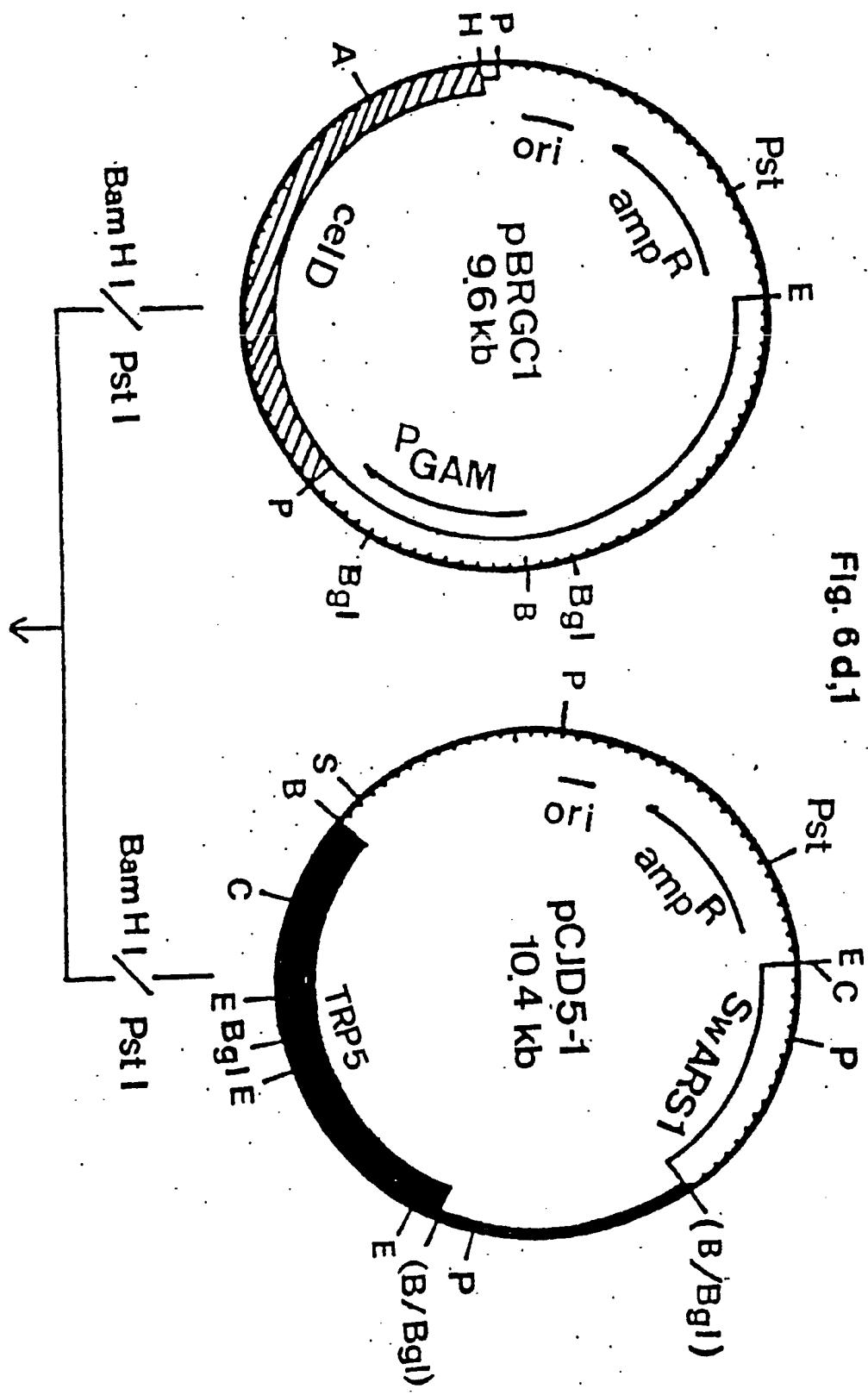


Fig.6c, 2



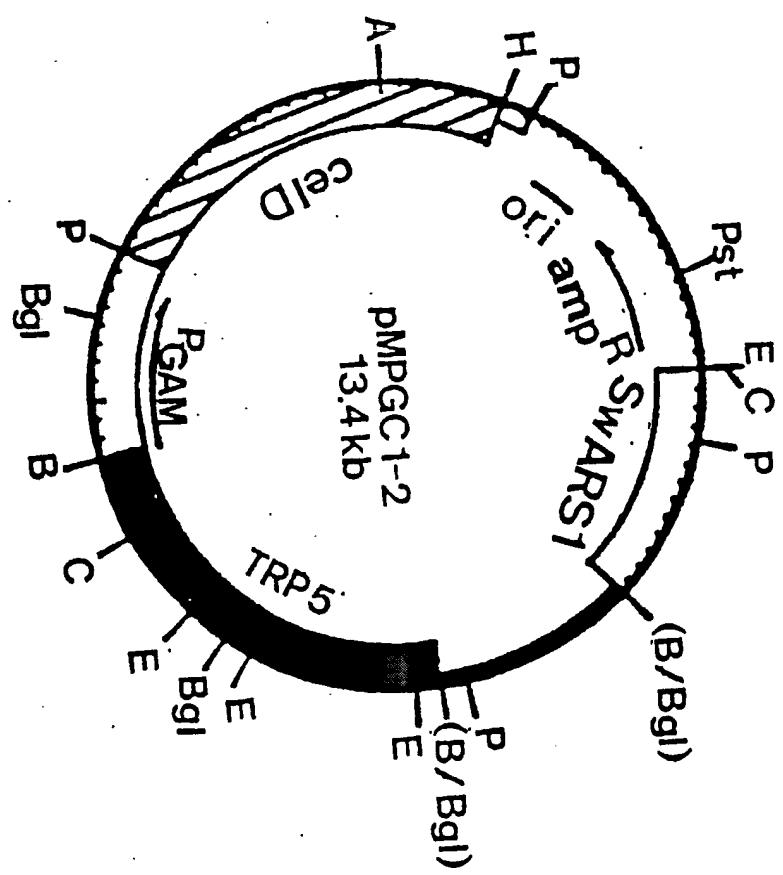


Fig.6d,2

Fig.7a

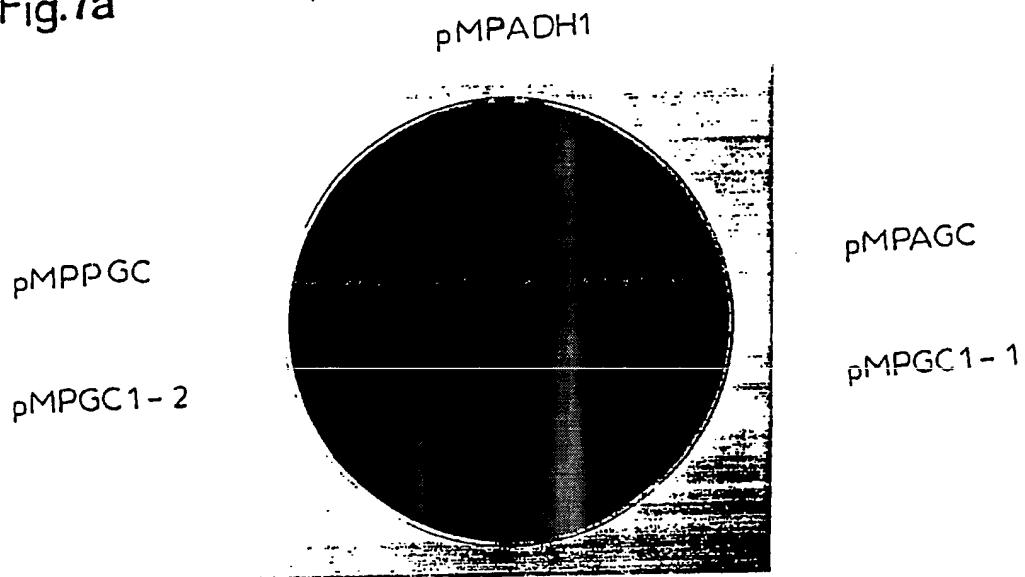


Fig.7b

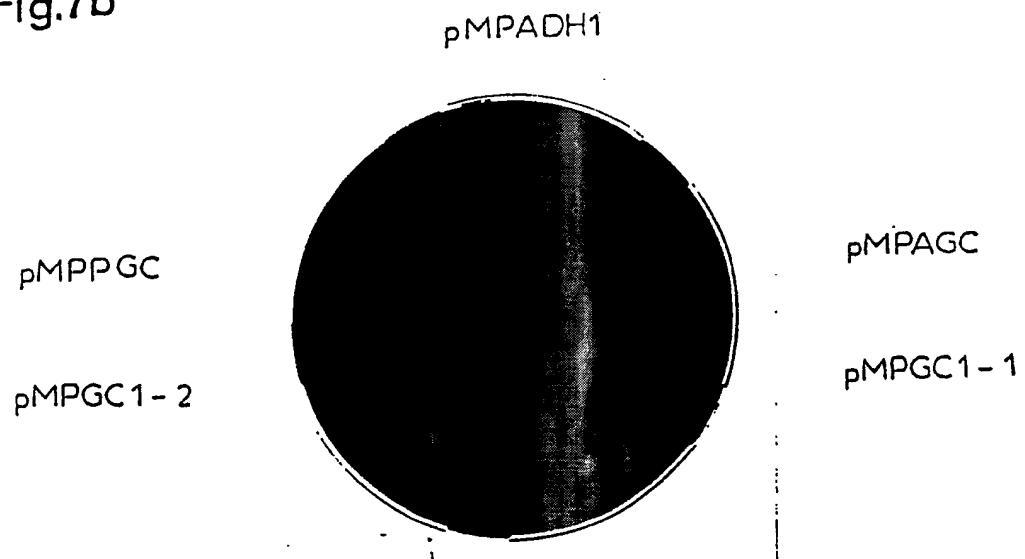
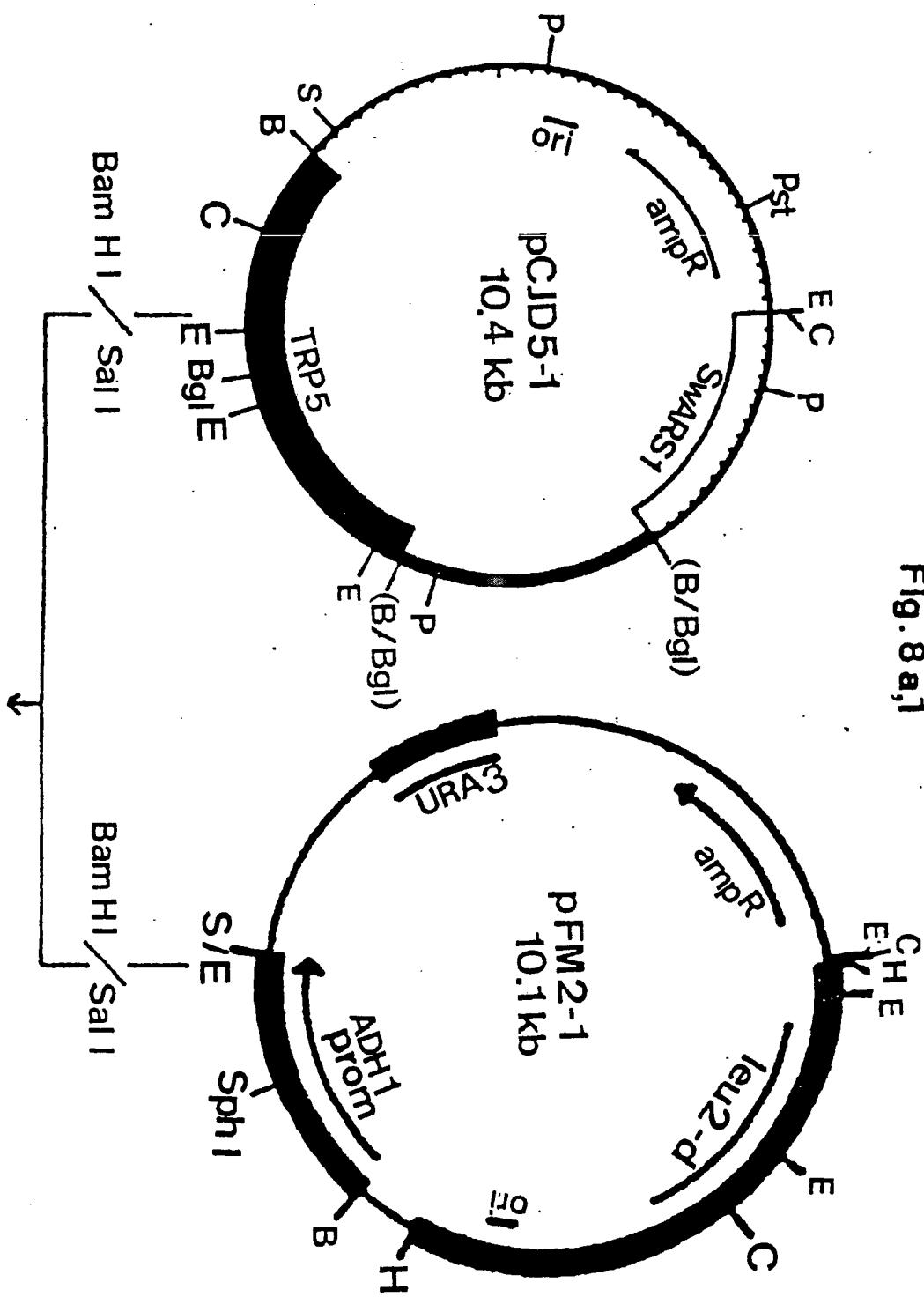


Fig. 8 a,1



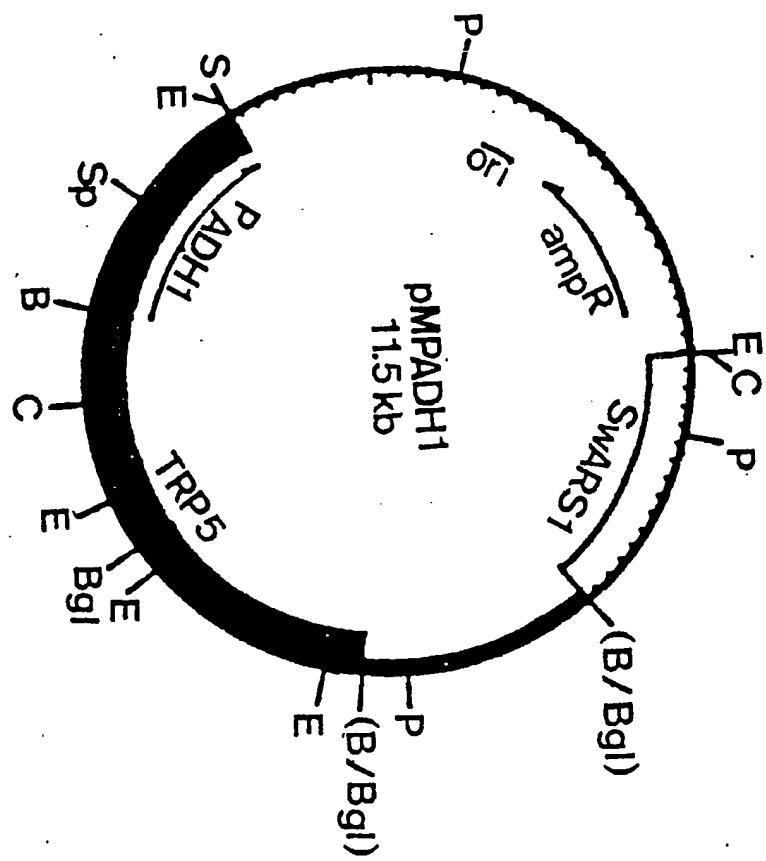
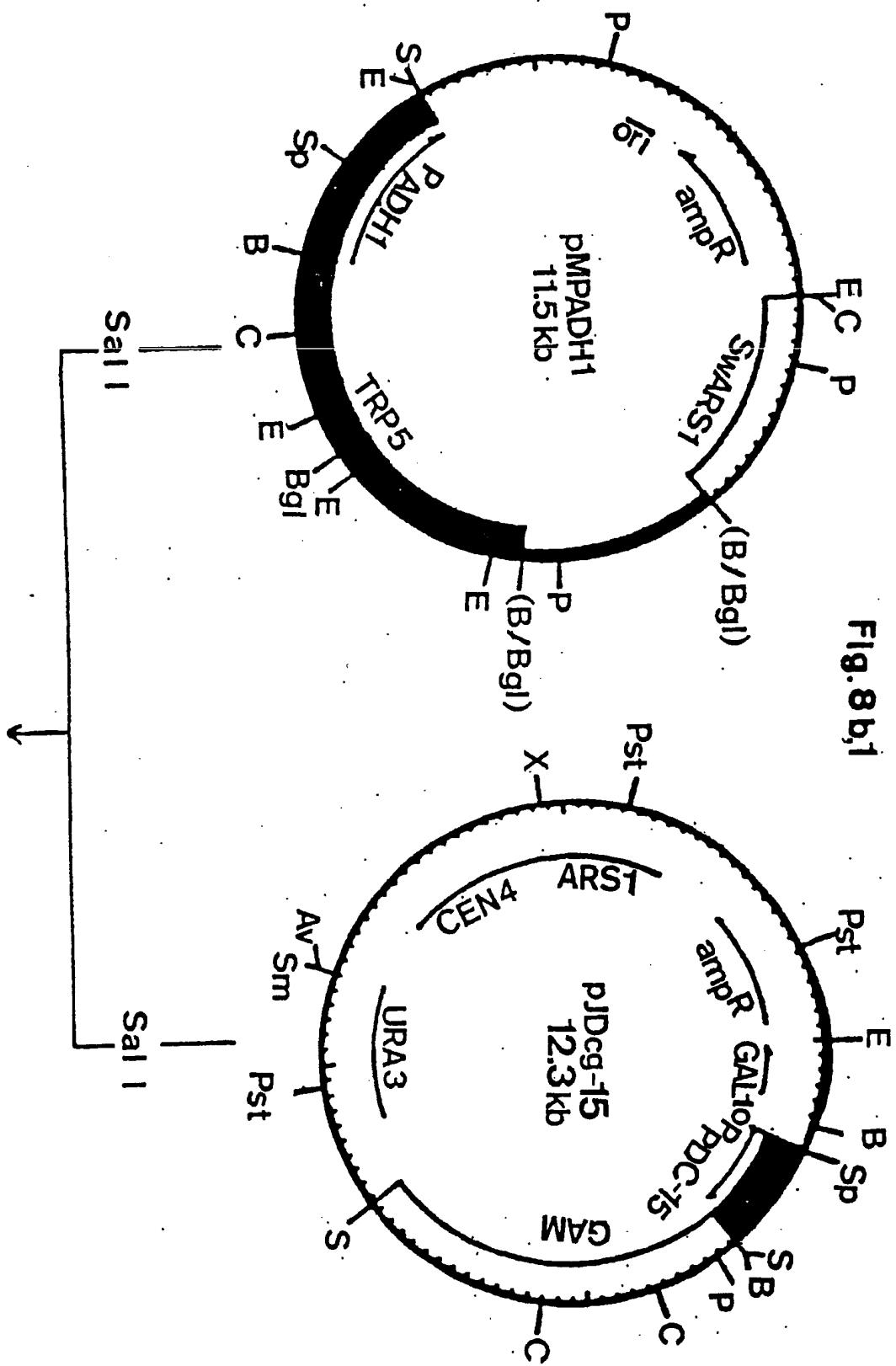


Fig.8a,2

Fig. 8 b,1



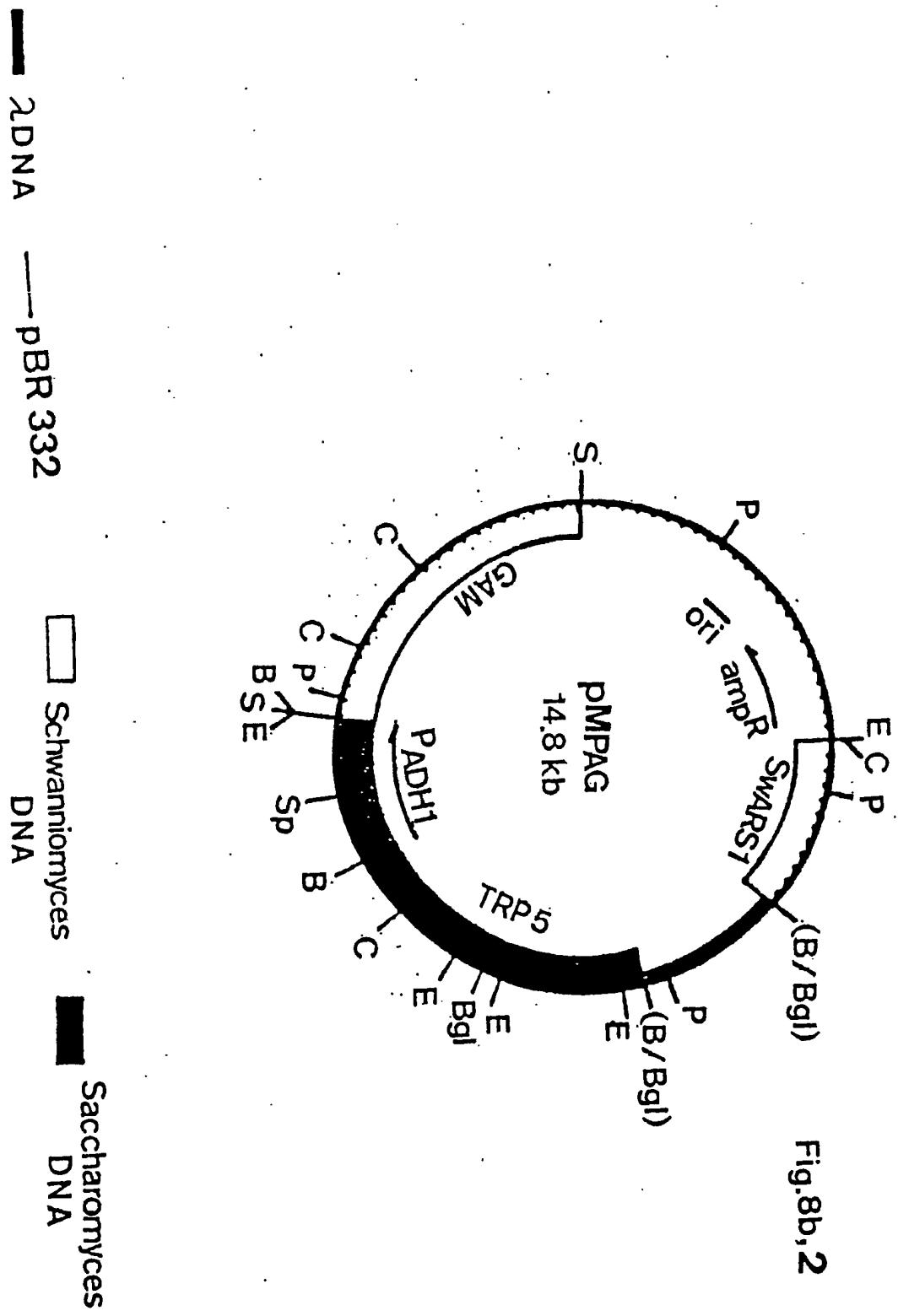


Fig.8b,2

Fig. 8 c,1

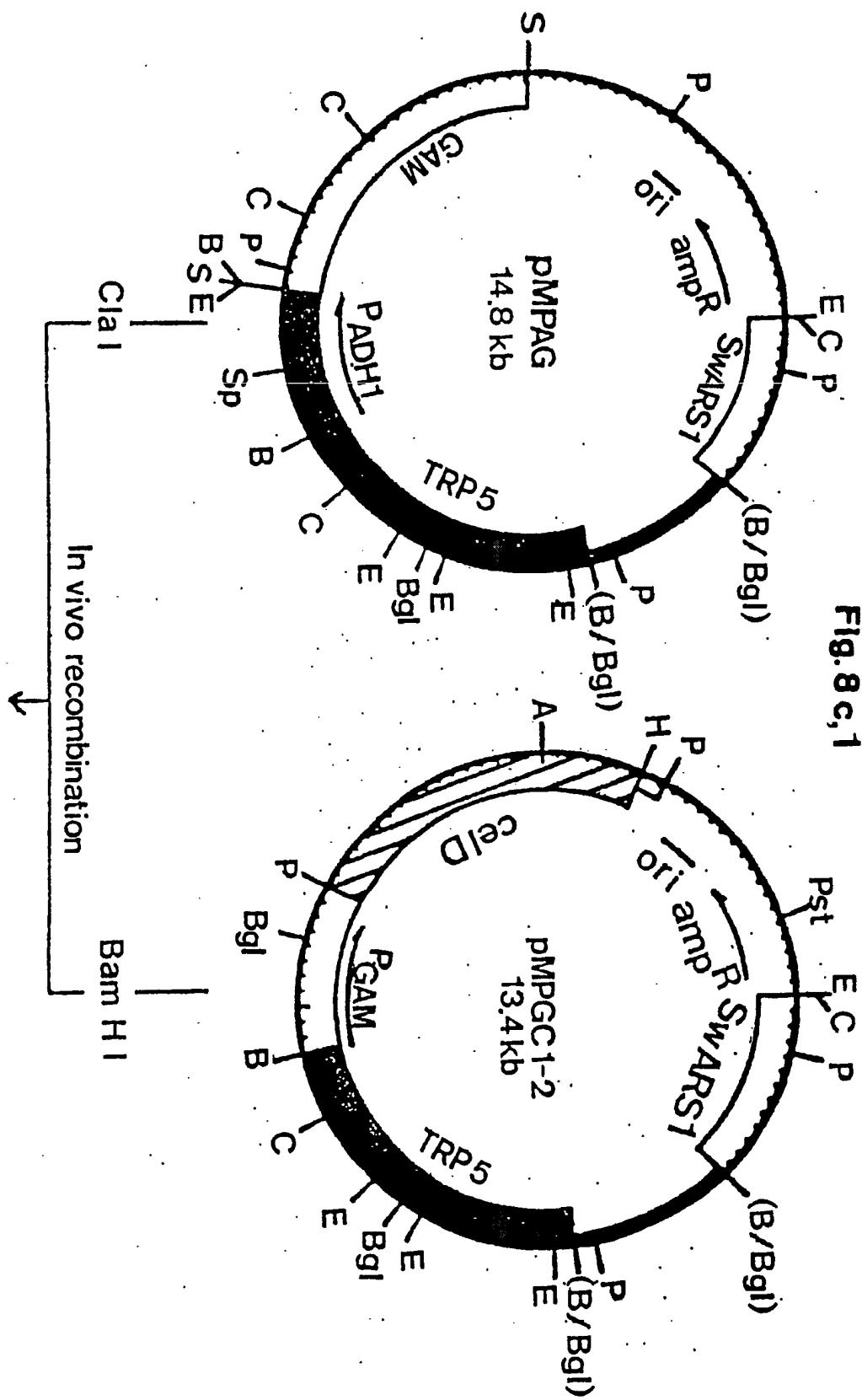


Fig.8c, 2

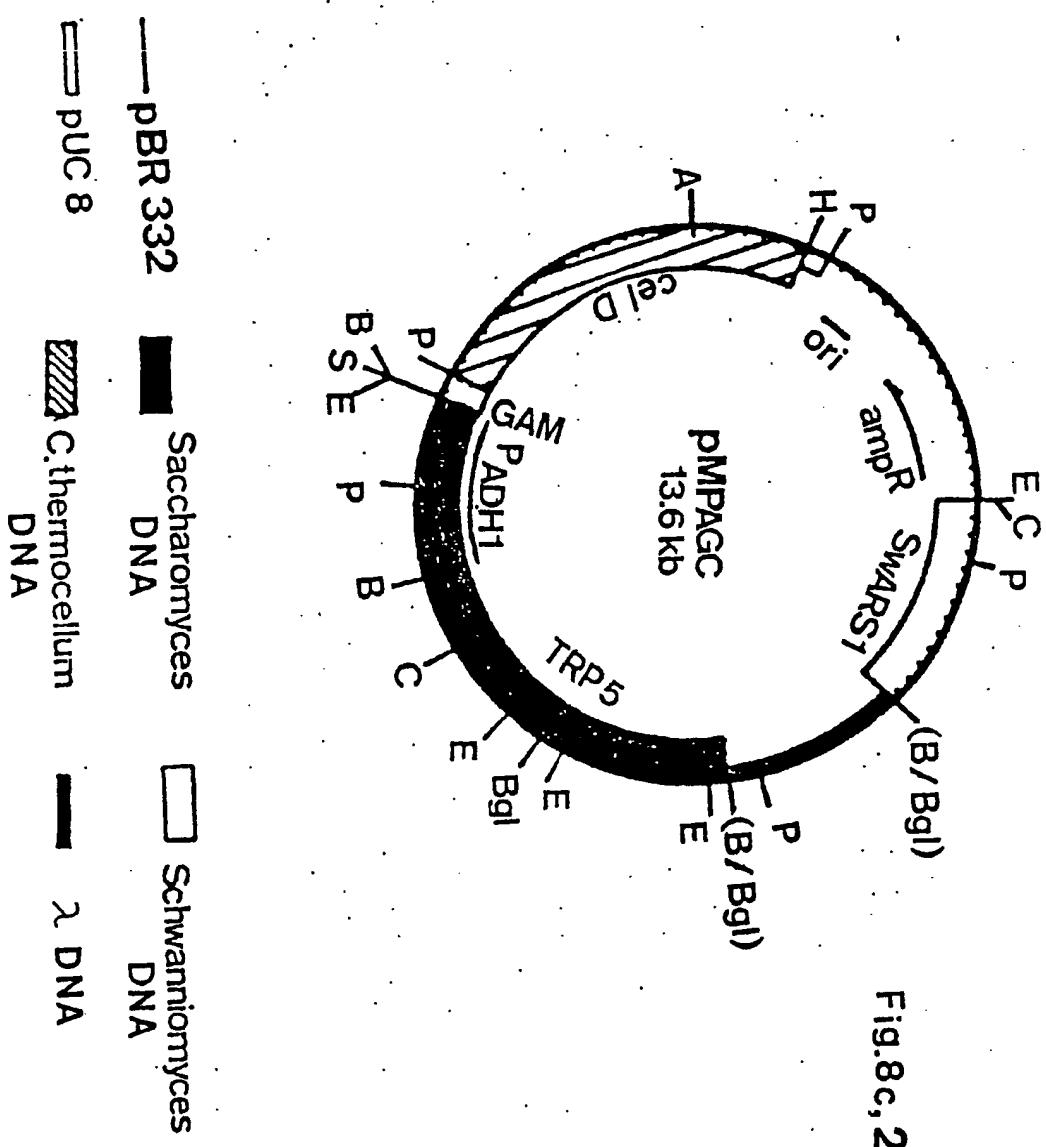
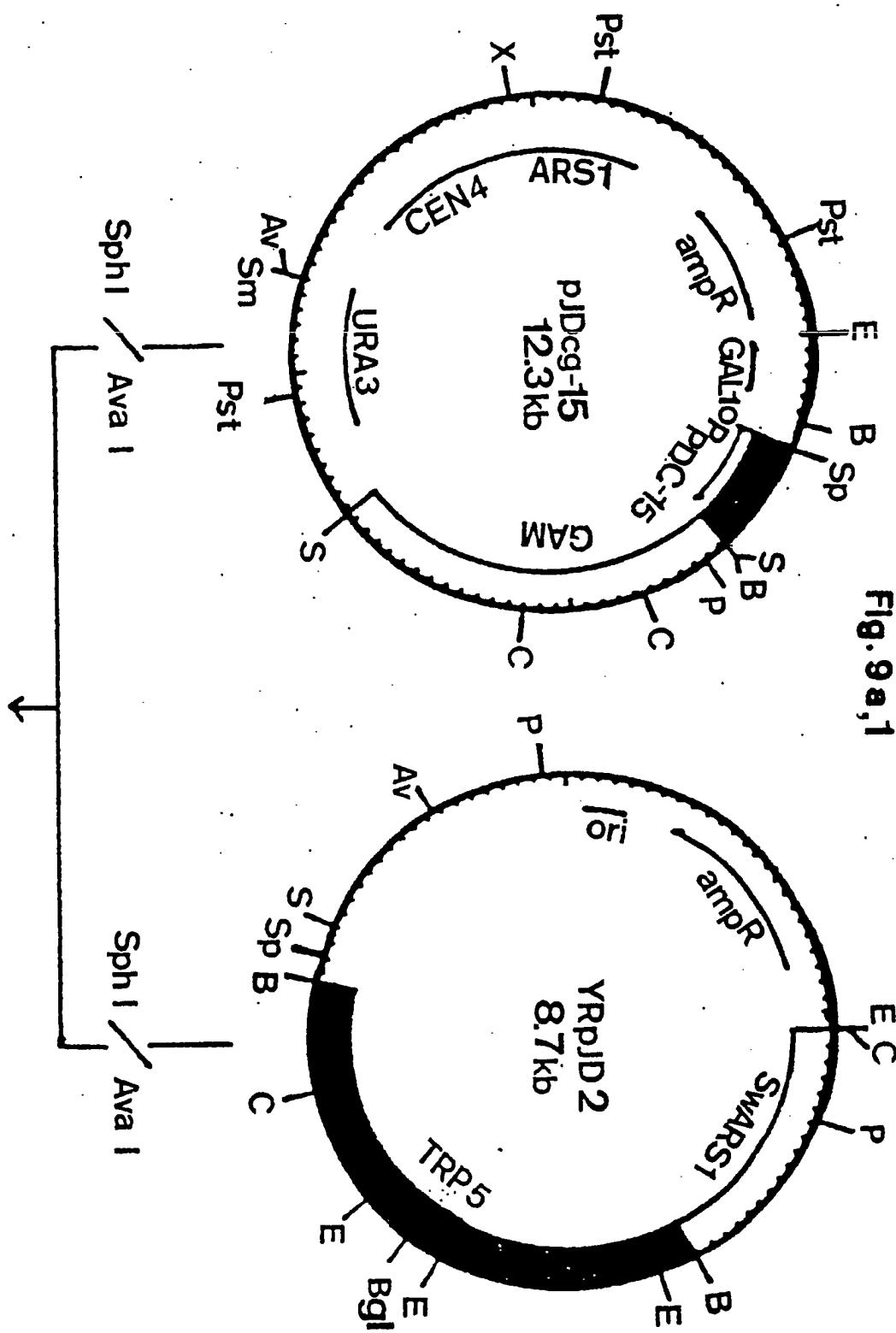


FIG. 88.8



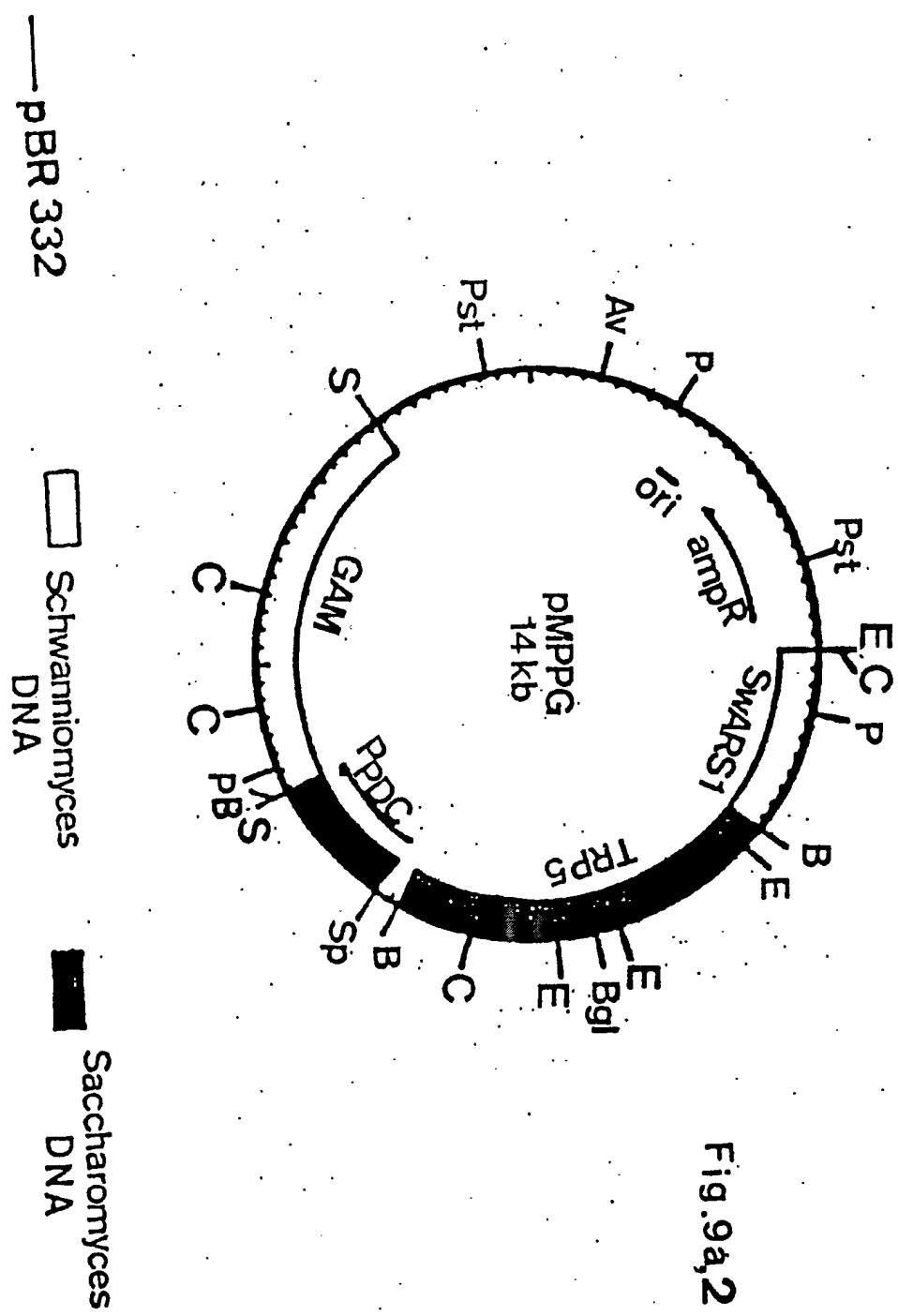


Fig. 9 b,1

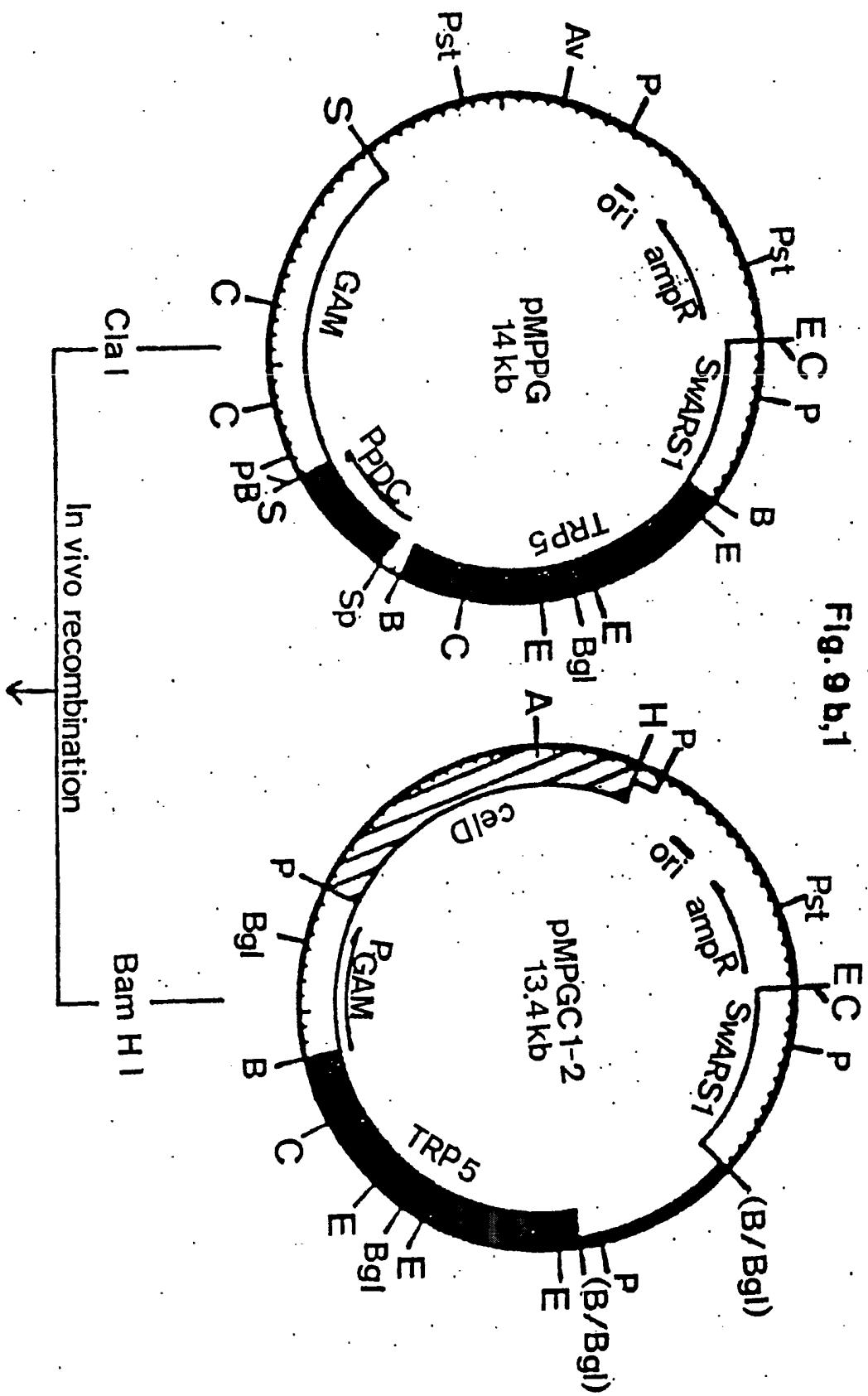


Fig.9b,2

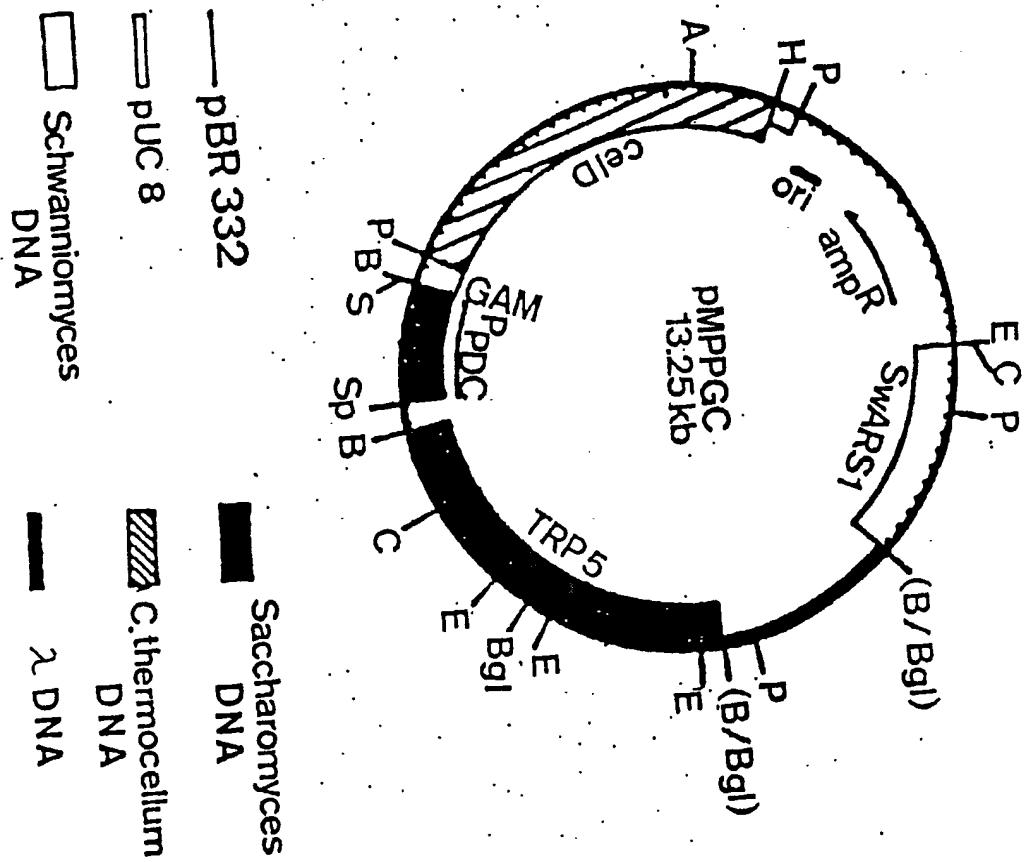


Fig. 10,1

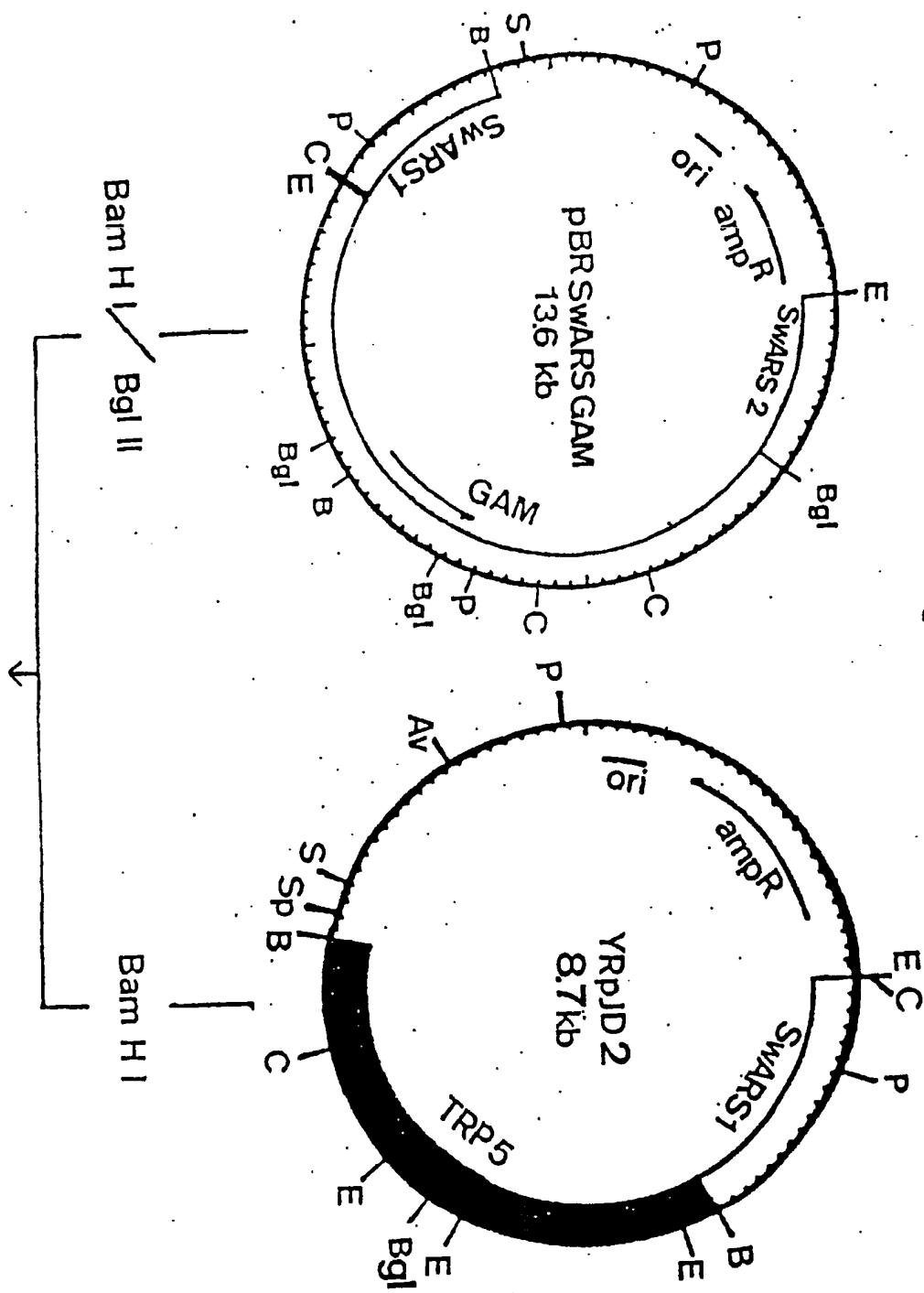
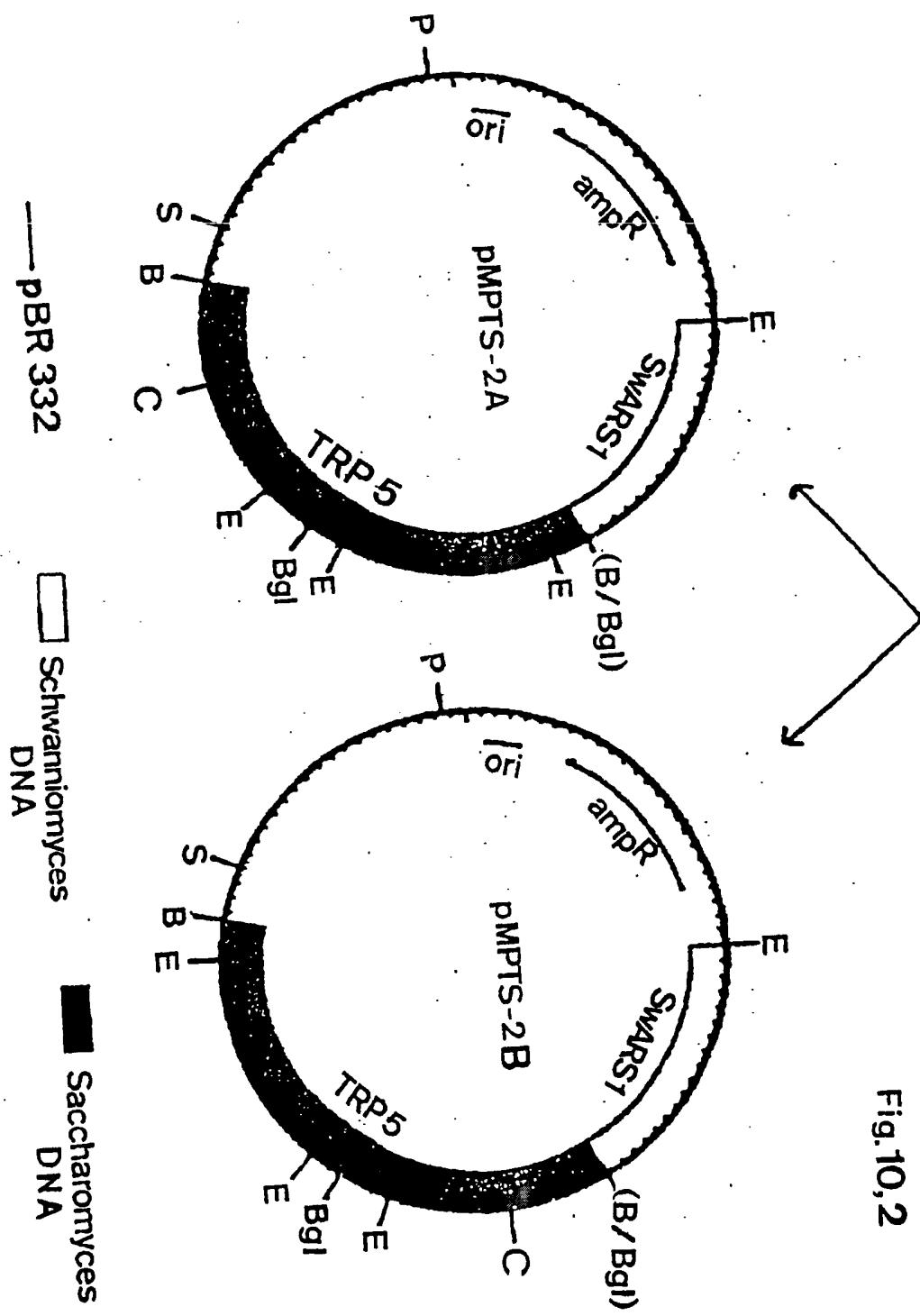


Fig.10,2





European Patent
Office

EUROPEAN SEARCH REPORT

Application Number

EP 89 10 7780

DOCUMENTS CONSIDERED TO BE RELEVANT															
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)												
A,D	EP-A-0 260 404 (HEINEKEN TECHNISCH BEHEER B.V.) ---	1	C 12 N 1/19 C 12 P 21/02 C 12 N 15/81 // (C 12 N 1/19 C 12 R 1:645)												
A,D	EP-A-0 257 115 (HEINEKEN TECHNISCH BEHEER B.V.) ---	1													
A	EP-A-0 256 421 (PHILLIPS PETROLEUM CO.) ---														
A	SCIENCE, vol. 229, 20th September 1985, pages 1219-1224, Washington, DC, US; R.A. SMITH et al.: "Heterologous protein secretion from yeast" ---														
A	CRC - CRITICAL REVIEWS IN BIOTECHNOLOGY, vol. 5, no. 2, 1987, pages 159-176, Boca Raton, Florida, US, W.M. INGLEDEW: "Schwanniomyces: a potential supereyeast?" -----														
			TECHNICAL FIELDS SEARCHED (Int. Cl.5)												
			C 12 N C 12 P												
<p>The present search report has been drawn up for all claims</p> <table border="1"> <tr> <td>Place of search</td> <td>Date of completion of the search</td> <td>Examiner</td> </tr> <tr> <td>THE HAGUE</td> <td>02-02-1990</td> <td>VAN PUTTEN A.J.</td> </tr> <tr> <td colspan="2">CATEGORY OF CITED DOCUMENTS</td> <td></td> </tr> <tr> <td colspan="2"> X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document </td> <td> T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document </td> </tr> </table>				Place of search	Date of completion of the search	Examiner	THE HAGUE	02-02-1990	VAN PUTTEN A.J.	CATEGORY OF CITED DOCUMENTS			X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document
Place of search	Date of completion of the search	Examiner													
THE HAGUE	02-02-1990	VAN PUTTEN A.J.													
CATEGORY OF CITED DOCUMENTS															
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document													



(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 394 538 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
16.10.1996 **Bulletin 1996/42**

(51) Int Cl.⁶: **C12N 1/19, C12P 21/02,**
C12N 15/81
// (C12N1/19, C12R1:645)

(21) Application number: **89107780.2**

(22) Date of filing: **28.04.1989**

(54) A yeast cell of the genus schwanniomyces

Hefezellen der Gattung-Schwanniomyces

Cellules de levure du genre schwanniomyces

(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE

(43) Date of publication of application:
31.10.1990 Bulletin 1990/44

(73) Proprietor: **RHEIN BIOTECH GESELLSCHAFT
FÜR NEUE BIOTECHNOLOGISCHE PROZESSE
UND PRODUKTE MBH
40595 Düsseldorf (DE)**

(72) Inventors:
• Hollenberg, Cornelius P., Prof. Dr.
D-4000 Düsseldorf (DE)
• Strasser, Alexander, Dr.
D-4000 Düsseldorf (DE)

(74) Representative: **Grünecker, Kinkeldey,
Stockmair & Schwanhäusser Anwaltssozietät
Maximilianstrasse 58
80538 München (DE)**

(56) References cited:
EP-A- 0 256 421 **EP-A- 0 257 115**
EP-A- 0 260 404

- SCIENCE, vol. 229, 20th September 1985, pages 1219-1224, Washington, DC, US; R. A. SMITH et al.: "Heterologous protein secretion from yeast"
- CRC - CRITICAL REVIEWS IN BIOTECHNOLOGY, vol. 5, no. 2, 1987, pages 159-176, Boca Raton, Florida, US; W.M. INGLEDEW.

EP 0 394 538 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description

The present invention relates to yeast cells of the genus *Schwanniomyces* and a process for the production of polypeptides.

For expression of eukaryotic proteins of biotechnological relevance there exists a variety of host systems, which are selected in compliance with the specific requirements. The most frequently used host organism is *Escherichia coli*, being the best known representative of prokaryotic host organisms; *Saccharomyces cerevisiae* and animal tissue cultures are used as eukaryotic host cells. For certain purposes, e.g. where expression of a glycoprotein is desired, expression in *E. coli* or other prokaryotic organisms is not appropriate, since same lack any glycosylation capability. On the other hand, the mass production of protein in tissue cultures still is cumbersome and expensive; therefore, the organism of choice in many instances is yeast of the well characterized genus *Saccharomyces*. *S. cerevisiae* is able to secrete efficiently small proteins, e.g. polypeptides like MG1 (Kurjan and Herskowitz, 1982), β -endorphin or α -interferon (Bitter et al., 1984), however proteins of higher molecular weight are captured in the periplasmic space (Emr et al, 1981) or the cytoplasma, thus causing decreased growth rates or even death of the cell.

Therefore it is an object of the present invention to provide an improved eukaryotic expression system, which is optionally capable of secreting proteins encoded by a foreign DNA into the culture medium.

This object has been solved by a yeast cell of the genus *Schwanniomyces*, wherein said yeast cell contains at least one expression cassette comprising:

- 20 a) a first DNA sequence serving as a regulon,
- b) optionally a second DNA sequence coding for a signal peptide
- c) a third DNA sequence coding for a foreign protein, and
- 25 d) optionally a fourth DNA sequence serving as a terminator, wherein the first, second and fourth DNA sequence are derived from yeast.

The invention, which comprises further subjects, is now described in a more detailed manner by the following description, examples and figures.

BRIEF DESCRIPTION OF THE FIGURES**Fig. 1**

35 A: Restriction map of the AMY1 gene encoding α -amylase and its 5' and 3' noncoding region

B: Nucleotide sequence of the AMY1 gene and its flanking region. In the 5' region the transcription initiation site is indicated by a dashed line above the sequence. Within the structural gene the presumptive signal peptide cleavage sites are indicated by star-symbols. The two potential glycosylation sites, the cystein residues presumably forming four disulfide bonds (concluded by homology to the *A. oryzae* enzyme) are underlined. Within the 3' end region the transcription termination signal TAG...TATGT...TTT is indicated by a solid line.

Fig. 2

45 A: Restriction map of the GAM1 gene encoding glucoamylase and its 5' and 3' noncoding region

B: Nucleotide sequence of the GAM1 gene and its flanking region. In the 5' region the transcription initiation site is indicated by a dashed line. Within the structural gene the presumptive signal peptides cleavage sites are indicated by star-symbols.

Fig. 3

55 M13 Xhol. The EcoRI-Scal fragment, containing a 1.8 kb promoter fragment and approx. 0.4 kb structural gene fragment of the α -amylase gene, was subcloned into M13mp8 previously digested with EcoRI and SmaI. By site directed mutagenesis a Xhol-site was inserted just in front of the ATG codon:

5'...TAAAATAAAAGCTCGAGATGAGATT...3'.

5

Positions where bases have been exchanged by site directed mutagenesis, are indicated by asterisks.

Fig. 4

10 pMaGAM-B/S. The BgIII fragment containing the glucoamylase gene and its 5' and 3' flanking region was sub-cloned into the BamHI site of vector pMAC5-8 (Kramer et al., 1984). By site directed mutagenesis using the gapped duplex DNA method (Kramer et al., 1984) a BamHI site and a Sall site were inserted just in front of the translation initiation codon ATG:

* *****

15

5'...CTCATGACTG TGTCGACGGATCCAAAGATGATTT...3'.

Positions affected by the site specific mutagenesis are indicated by asterisks.

20

Fig. 5

YRp7 α GAM. The GAM gene was isolated from pMaGAM-B/S as a Sall fragment and ligated into the XhoI site of M13 α Xho (Fig. 3). The BgIII fragment harbouring the α -amylase promoter GAM gene fusion was ligated into the single BamHI site of the *S. cerevisiae* vector YRp7 resulting in plasmid YRp7 α GAM.

25

Fig. 6

Construction of *S. occidentalis* vectors for expression of foreign genes using the promoter and the signal sequence of the glucoamylase gene.

30

a: Insertion of the 4.0 kb EcoRI-PvuII fragment from pBRSwARSGAM, containing the complete GAM promoter and the first 208 bp of the GAM encoding sequence into pRR322, cleaved with EcoRI/PvuII, resulting in vector pBRGAM.

35

b: Insertion of the 3.4 kb PvuII fragment from plasmid pCT603, containing the structural gene coding for cel-lulase from base pair position +112, into the PvuII site of plasmid pBRGAM, resulting in vector pBRGC1.

40

c: Insertion of the 5.5 kb BgIII-PstI fragment from plasmid pBRGC1 into BamHI/PstI cut pCJD5-1. The resulting plasmid, pMPGC1-1 contains 321 bp of the GAM 5' non coding region and the first 208 bp of the GAM coding region fused to position +112 of the celD gene.

45

d: Insertion of the 6.5 kb BamHI-PstI fragment from plasmid pBRGC1 into BamHI/PstI cut pCJD5-1. The resulting plasmid, pMPGC1-2 contains 1.3 kb of the GAM 5' non coding region and the first 208 bp of the GAM coding region fused to position +112 of the celD.

45

Fig. 7

Promoter studies using the qualitative endoglucanase D activity assay (congo red staining) with different transformants: NGA58: pMPAGC, NGA58: pMPPGC, NGA58: pMPGC1-1, NGA58: pMPGC1-2

50

a: Growth on 2% maltose

b: Growth on 4% glucose

Fig. 8

Construction of a *S. occidentalis* vector for expression of foreign genes under control of the ADH1 promoter.

55

a: Insertion of the 1.45 kb BamHI-Sall ADH1 promoter fragment of plasmid pFM2-1 into pCJD5-1, previously cut with BamHI/Sall resulting in plasmid pMPADH1.

b: Insertion of the 3.2 kb Sall fragment of pJDcg-15 into the Sall site of pMPADH1, resulting in pMPAG.

c: Exchange of the GAM promoter in plasmid pMPGC1-2 (Fig 6d) for the ADH1 promoter from plasmid pMPAG by *in vivo* recombination, resulting in plasmid pMPAGC.

5 Fig. 9
Construction of a *S. occidentalis* vector for the expression of foreign genes under control of the PDC promoter.

a: Insertion of the 6.0 kb SphI-Aval fragment from plasmid pJDcg-15, containing the PDC promoter fused to the GAM gene into YRpJD2 previously cleaved with SphI-Aval, resulting in plasmid pMPPG.

10 b: Exchange of the GAM promoter in plasmid pMPGC1-2 (Fig. 6d) for the PDC promoter from plasmid pMPPG by *in vivo* recombination, resulting in plasmid pMPPGC.

Fig. 10

15 Construction of an autonomous replicating vector (containing SwARS2) for transformation of *S. occidentalis*. The 3.2 kb TRP5 BamHI fragment from plasmid YRpJD2 was ligated with the 5.5 kb BamHI-BglII fragment of plasmid pBRSwARSGAM (Fig 6a), resulting in pMPTS2-A and pMPTS2-B.

ABBREVIATIONS USED THROUGHOUT THE APPLICATION

20 Restriction endonucleases :

Aa	-AatII
A	-Asp718
25 B	-BamHI
Bg	-BglII
C	-ClaI
E	-EcoRI
EV	-EcoRV
30 H	-HindIII
P	-PvuII
Ps	-PstI
S	-Sall
Sc	-ScaI
35 Sm	-SmaI
Sp	-SphI
X	-XbaI

40 Throughout this application various publications are referenced by mention of the first author and the year of publication within parentheses. Full citation of the references may be found at the end of the Specification as an annex, listed according to their alphabetical order immediately following the specification. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art as known to those skilled therein as of the date of the invention described and claimed herein.

45 The yeast cell according to the present invention is characterized by containing an expression cassette which in turn is composed of three to four different components, being arranged in operable linkage in order to ensure efficient expression of a desired polypeptide, optionally followed by secretion of same. These components are defined to be the following :

50 a) A first DNA sequence serving as a regulon. The term "a regulon" as used throughout the application comprises any cis-acting DNA sequence involved in the regulation of expression of a given structural gene. The term thus embraces sequences preceding a given coding sequence, for example, a promoter and sequences which are recognised by transcription factors or other proteins involved in transcription regulation. The "first DNA sequence" is thus not limited to sequences corresponding to known promoters. Also included are DNA sequences corresponding to natural occurring regulons, which have undergone insertion, deletion or substitution events, but still retain their activity as regulon. Further the term "regulon" includes DNA sequences involved in regulation of transcription, which have been synthesised chemically or obtained by methods of gene technology.

55 b) An optional second DNA sequence coding for a signal peptide. The meaning of this second DNA sequence

includes various DNA sequences, directing the secretion of any polypeptide containing the thus encoded peptide at its N-terminal end.

5 c) A third DNA sequence coding for a foreign protein. This coding sequence coding for the signal peptide in order to retain a proper reading frame for the third DNA sequence. The fusion thus requires application of methods or recombination DNA technology known to every person skilled in the art. For methods enabling to design proper fusions it is in general referred to Maniatis et al. (1982), wherein appropriate methods, e.g. various ways to treat restriction fragment ends, cloning of linker molecules and similar procedures are described in great detail. The foreign protein coded for by the third DNA sequences may be any protein which to express there exists some need. Since the secretion system is capable of secretion not only of small molecules, but also of large proteins, expression of any desired protein should be possible. For examples it is referred to the discussion below.

10 15 d) Optionally a fourth DNA sequence serving as a terminator. As the terminator any sequence may be used, which efficiently terminates transcription in the respective host organism. Terminators may be for example sequences prone to formation of hairpin structures and/or polyadenylation sites. In a preferred embodiment the terminator is derived from the same gene as the second DNA sequence, i.e. the regulon, is derived from.

The four components are combined with each other using known techniques. Optionally the proper fusion will be verified for example by sequencing the boundaries.

20 The yeast cell according to the present invention contains at least one expression cassette comprising some or all of the above mentioned DNA sequences. It provides a convenient tool for the industrial expression of any desired protein.

25 In a preferred embodiment said first and/or second and/or fourth DNA sequence are derived from a gene coding for an amylolytic enzyme, for example α -amylase or glucoamylase, which enzymes in most organisms are induced upon contact with starch. The DNA sequences may be derived from a yeast of a member of one of the genera Debaromyces, Saccharomyces, Lipomyces, Pichia, or Saccharomyces, but yeast of the genus Schwanniomyces are preferred.

30 Yeast species of the genus Schwanniomyces have an efficient expression and secretion apparatus, enabling them to grow on starch as a sole carbon source. Starch is hydrolysed to glucose by two amylolytic enzymes, secreted into the supernatant, namely α -amylase (α -1,4-glucan 4-glucanohydrolase E.C.3.2.1.1) and glucoamylase (syn. amyloglucosidase) (α -1,4-glucan glucohydrolase, E.C.3.2.1.3.; with debranching activity E.C.3.2.1.9.). The amylolytic enzymes of Schwanniomyces occidentalis, formerly also called Schwanniomyces castellii and Schwanniomyces alluvius (Price et al., 1978), are well documented (Oteng-Gyang et al., 1981; Sills et al., 1982, 1984a, 1984b; Wilson et al., 1982).

35 The genes for α -amylase and glucoamylase from a yeast of the genus Schwanniomyces have been disclosed by the present inventors in EP-A-0260404. However, at that time the DNA or peptide sequences directing secretion have not been identified. Furthermore their use for the establishment of an efficient expression and secretion system in yeast was not known then.

40 It is preferred, to use as a first DNA sequence part or all of a 1.8 kb BgIII-Xhol fragment which precedes the structural gene coding for the Schwanniomyces α -amylase, or part or all of a 1.3 kb BamHI-PvuII fragment preceding the structural gene coding for the Schwanniomyces occidentalis glucoamylase. Further if only smaller fragments are to be used, it is preferred, to use a DNA sequence comprising part or all of the DNA sequence corresponding to bases -1 to -540 of the region preceding the structural gene coding for α -amylase or part or all of the DNA sequence corresponding to bases -1 to -320 of the region preceding the structural gene coding for glucoamylase. The respective DNA fragments are fully active as promoters, inducible by starch, dextrin, maltose.

45 It is obvious, that replacement of bases or base pairs, not involved in the transcriptional regulation, still are within the spirits of the present invention. Thus natural mutations or synthetic equivalents also are embraced by the present invention.

50 In yeast cells of the genus Schwanniomyces it is also possible to use promoters derived from yeast genes obtained from many other yeast genera. Preferred are regulons involved in the expression of ADH1, ADH2, PDC1, GAL1/10, PGK and GAPDH, or LAC4 which are obtained from Saccharomyces cerevisiae, Kluyveromyces Lactis or Schwanniomyces occidentalis.

55 Furthermore, it is possible to use viral regulons and/or terminators, for example regulons and terminators obtained from E.coli T-phages. The phages possess extraordinarily strong functional sequences for performance of their lytic functions. In a preferred embodiment sequences of phage T7 are used, which are regulated by expression of T7 RNA polymerase.

In order to increase the over all expression level and furthermore, in order to facilitate purification of produced polypeptides, it is most desirable, that the polypeptides are secreted into the culture medium. The present invention provides signal sequences having the invaluable advantage of being derived from genes of amylolytic and thus induc-

ible enzymes, which upon induction have to be secreted very efficiently. A further advantage is, that induction of amylolytic enzymes does not require complicated media, lacking for example a specific amino acid or phosphate, but only require the culture to be transferred to a starch containing medium. Signal sequences, which are preferably used for secretion of large proteins include part or all of at least one of the following peptides:

5

a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle
LeuAlaPheSerArgLeuValSerAla;

10

b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle
LeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;

15

c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAla;

20

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAlaIleGlnAlaAlaProAla;

30

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAlaIleGlnAlaAlaProAlaSerSerIleGlySerSerAlaSerAla

35

In a further embodiment the above mentioned peptides are coded for by DNA sequences, which correspond to all or part of the following DNA sequences:

40

a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATTT
AGCATTCTCAAGATTGGTATCTGCT;

45

b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TAGCATTCTCAAGATTGGTATCTGCTCAACCGATTATTTGACATGAGA;

55

c) ATGATTTCTGAAGCTGATTAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCT;

d) ATGATTTTCTGAAGCTGATTAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCC;

5

e) ATGATTTCTGAAGCTGATTAAGTATAGTAATTGGTTGGATTAGTTAGTGCTATCCAAGCAGCCCCCTGCCTCTCGATTGGATCTAGTGCTTCAGCA.

10

15 It will be understood, that the replacement of amino acids by residues having comparable properties and use of alternative codons in compliance with the possibilities provided by the genetic code will not have any impact on the yeast cell according to the present invention.

15 yeast cell according to the present invention.

The polypeptide to be expressed in a yeast cell according to the present invention is coded for by a third DNA sequence. Said third DNA sequence may comprise natural or synthetic DNA, even DNA containing intervening sequences. Examples for proteins being expressed by the yeast cell according to the present invention comprise proteins of enormous importance for scientific or medical purposes, for example cellulase, interleukine, interferon, insulin like 20 growth factor, lymphokine, human growth factor, nerve growth factor, aprotinin, insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines or human granulocyte/macrophage colony stimulating factor.

The yeast cell according to the present invention may optionally comprise a terminator as component of said expression cassette. Preferred terminators are derived from genes of one of the following genera:

25 expression cassette. Preferred terminators are derived from *Saccharomyces*, *Pichia*, *Hansenula*, but preferably *Schwanniomyces*. Examples of preferred terminators are those derived from *Schwanniomyces* α -amylase gene or *Schwanniomyces* glucoamylase gene, part or all of which may be used. Such terminators are comprised by nucleotides 1537 to 1740 of the *Schwanniomyces* α -amylase gene or nucleotides 2875 to 3320 of the *Schwanniomyces* glucoamylase gene.

30 It is obvious, that in case of any of the DNA sequences mentioned modifications, which do not impair the biological function of the respective DNA fragment still are covered by the present application. There are numerous possibilities to change the DNA sequences mentioned above slightly without impairing the biological function of said fragment gene.

35 The yeast cell according to the present invention, carrying an expression cassette composed of the above discussed components, may be carried by a circular or linear vector, which optionally comprises one or more selective marker genes. The present inventors demonstrated, that a lot of marker genes so far used for transformation of *Saccharomyces*, for example TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, which may be obtained from *Saccharomyces*, *Hansenula*, *Pichia* or *Schwanniomyces*, are functional in *Schwanniomyces occidentalis*. Furthermore it is also possible to use the amylolytic properties of α -amylase or glucoamylase in order to select transformed organisms. The corresponding genes have been isolated and characterized. AMY1 and GAM1 are preferably obtained from *Schwanniomyces*.

40 The yeast cell according to the present invention may contain the expression cassette as an insert into one of the
endogenous yeast chromosomes. Furthermore, there is however the possibility, that a DNA sequence capable of con-
trolling autonomous replication and even also stable maintenance of the vector in the host organism, is introduced into
the yeast cell. Said DNA sequences are the so called ARS sequences, capable of controlling autonomous replication
and stable maintenance in the respective host organism. Said DNA sequences are preferably derived from Saccha-
45 romyces, Pichia, Hansenula or Kluyveromyces, most preferred from Schwanniomyces.
46 ABC-1, 151934/03

In order to ensure effective replication in the prospective host organism, it is preferred to use an ARS, derived from the respective host organism.

The present invention provides one new SwARS sequence. The DNA sequence SwARS2 is capable of controlling autonomous replication and stable maintenance of the vector in a *Schwanniomyces* yeast cell. The detailed sequence of SwARS has not yet been determined. SwARS2 is localised in the *Schwanniomyces* genome in the 3' region of the glucoamylase gene, as shown for example in Fig. 10. SwARS2 is easy to identify and efficiently allows autonomous replication of any plasmid, exhibiting this sequence. The so far known *Schwanniomyces* autonomously replicating sequences, SwARS1 and SwARS2 are comprised on a *Clal-Bam*HI fragment (SwARS1) adjacent to the *Eco*RI-fragment shown in Fig 2a (see Fig. 10) and/or an *Eco*RI-*Bgl*III fragment contained in the *Eco*RI fragment shown in Fig 2a (SwARS2).

For the SwARS sequences it is also not necessary, in any case to use the sequence as it occurs naturally. DNA sequences, that have been obtained by any kind of modifying treatment, for example by deletion, insertion or substitution of one or more nucleotides or base pairs, are still within the scope of the present invention, provided the modi-

fications retain or improve the biological function of the yeast cell according to the present invention.

In a preferred embodiment the yeast cell according to the present invention comprises a DNA sequence, which is homologous to genomic *Schwanniomyces* DNA. Provision of structures of homologous DNA enable the DNA fragment flanked by such homologous sequences to be inserted by homologous recombination. It is thus preferred to have flanked by such homologous sequences, which do not represent any essential gene of the *Schwanniomyces* genomic DNA, but rather any gene, whose expression product may be supplemented by additives to the respective medium or were lack of the respective expression product does not have any serious impact on the transformed yeast cell.

5 If no SwARS sequence or any equivalent sequence, for example originating from any other yeast species, is available, the expression cassette can be used without an autonomously replicating agent. In this case the expression cassette tends to integrate into the genome, for example via homologous recombination. Preferably there are several 10 copies inserted into one or more chromosomes.

In this case it is preferred, to have multiple copies of the expression cassette inserted, in order to ensure simultaneous transcription and translation of several copies.

15 When the yeast cell according to the present invention is derived from a yeast of the genus *Schwanniomyces* the polypeptide products observed following expression of the respective foreign gene in said yeast cells exhibit glycosylation patterns, which closely resemble or even match those of the desired eucaryotic protein. Yeast cells of the genus *Schwanniomyces* are thus preferred for expression of eucaryotic proteins.

20 The above discussed yeast cell according to the present invention can be used in a process for the production of a polypeptide, wherein a yeast host organism is cultivated under suitable conditions and the polypeptide is recovered in a manner known *per se*. The conditions of cultivation depend on the respective host organism used. Examples for 25 cultivation conditions for yeast of the genus *Schwanniomyces* are provided in the examples below. The polypeptide, produced by the yeast according to the present invention, is recovered by conventional purification methods, which depend on the properties of the expressed protein, the availability of specific absorbing agents and so on.

25 In a preferred embodiment yeast cells of the species *Schwanniomyces occidentalis* are used as a host organism and are cultivated on starch, dextrin, maltose and/or plant biomass as the respective carbon source. As commonly known, these carbon sources are available at low costs.

Furthermore, there is a possibility to use cells of the genus *Schwanniomyces* for the production of single cell protein. The single cell protein may then be subjected to any further use as it is required.

30 In a preferred embodiment the polypeptide produced in a process according to the present invention is an amylolytic enzyme, for example an α -amylase or glucoamylase. These enzymes are commercially used to debranch and degrade starch, which thereupon may be further processed, for example in brewery or baking processes. There is a high need for the provision of starch degrading enzymes.

35 In a further embodiment of the process according to the present invention production of a polypeptide may be achieved by providing expression cassettes as discussed above containing a yeast α -amylase or glucoamylase signal sequence and inserting the respective expression cassette into a yeast of one of the genera *Saccharomyces*, *Kluyveromyces*, *Hansenula*, *Pichia* or *Schizosaccharomyces*. Production and secretion of foreign proteins in the respective yeast host organism is obtained by applying the appropriate inductive measure, whereupon the produced protein is secreted due to the presence of the yeast α -amylase or glucoamylase signal sequence. Since it is known from experiments performed by the present inventor, that these signal sequences are recognised in virtually any yeast genus, 40 the α -amylase or glucoamylase signal sequences may be applied in many different yeast with success. Also in this case, it is preferred, to use the α -amylase or glucoamylase signal sequences derived from the respective genes of *Schwanniomyces occidentalis*.

45 In summary the yeast cells according to the present invention are well suited for transformation and expression of any desired foreign gene. As already outlined above, expression of prokaryotic as well as eucaryotic genes is possible. However, for expression of eucaryotic proteins yeast cells of the genus *Schwanniomyces* containing some or all of the DNA sequences mentioned above are the most preferred yeast cells.

CHARACTERISATION OF THE α -AMYLASE AND GLUCOAMYLASE GENE EXPRESSION IN DIFFERENT YEAST GENERA:

50 The N-termini of the deduced amino acid sequences of the α -amylase gene (Fig. 1B) and glucoamylase gene (Fig. 2B) exhibit the features of typical leader sequences of secretory proteins (Von Heijne 1983, Perlman and Halvorsson 1983). The N-terminal amino acids of mature α -amylase secreted by *S. occidentalis* are determined to be Asp/Valser. This indicates processing between Arg33 and Asp31 which is unusual and not according to the Von Heijne's model of site specificity for signal peptidase cleavage. This putative cleavage site was confirmed also for α -amylase expressed in *S. cerevisiae*, *Pichia stipidis*, *K. lactis*, *S. pombe* and *H. polymorpha*. However, in these yeast species processing is also found to occur between Ala 25 and Gln 26, resulting in the N-terminal sequence GlnProle.

α -amylase isolated from a culture of *S. occidentalis* migrates as a single band of 55 kDa on SDS polyacrylamide

gels: following Endoglucosidase H treatment and subsequent gel electrophoreses a single band of 54 kDa is observed, indicating that only one of the two potential N-glycosylation sites of the protein is used for glycosylation. It is noteworthy that high mannose glycosylation, as very often found by proteins secreted in *S. cerevisiae* is not accompanying secretion in *S. occidentalis*.

5

IDENTIFICATION OF THE TRANSCRIPTION INITIATION SITES AND PROMOTER FUNCTION OF THE α -AMYLASE AND GLUCOAMYLASE GENE.

10

By S1 mapping of polyA-mRNA (Sharp et al., 1980) the transcription initiation site of the α -amylase gene could be localised to be in the region of -42 to -30 (indicated by an interrupted line above the relevant sequence in Fig.1B) and the transcription initiation site of the glucoamylase gene was determined to be in the region -7 to -10 (underlined in Fig.2B).

15

CONSTRUCTION OF AN α -AMYLASE - GLUCOAMYLASE HYBRID GENE

20

In order to facilitate use of the α -amylase promoter and glucoamylase structural gene in further constructions additional restriction sites were inserted in front of the translation initiation codon of both genes using oligonucleotide directed mutagenesis as described in the AMERSHAM protocol. Correct mutagenesis was verified by sequencing according to Sanger (1977). The boundaries between the original *Schwanniomyces* genes and the inserted polylinker are shown below:

α -amylase :

25

5' ...TAAAATAAAAGCTCGAGATGAGATT...3'
XhoI

30

Glucamylase :

35

5' ...CTCATGACTGTGTCGACGGATCCAAGATGATTT...3'
SalI BamHI

40

The manipulations resulted in plasmids M13 α Xho (Fig.3) and pMaGAM-B/S (Fig.4).

Thus, both the α -amylase promoter and the structural glucoamylase gene can be isolated from plasmid M13 α Xho and pMaGAM-B/S as suitable DNA fragments. The GAM gene was isolated from pMaGAM B/S as a 3.2 kb Sall fragment and ligated into the XhoI site of M13 α Xho. The BgIII fragment harbouring the α -amylase promoter/GAM gene fusion was ligated into the single BamHI site of the *S. cerevisiae* vector YRp7 (Struhl et al., 1979) resulting in plasmid YRp7 α GAM (Fig 5). *S. cerevisiae* strain YNN27 was transformed with this plasmid selecting for TRP prototrophy. Transformants secreted active glucoamylase into the culture supernatant (50 mU/ml) indicating that the α -amylase promoter fragment can direct expression and secretion of foreign proteins.

50

SECRETION OF α -AMYLASE AND GLUCOAMYLASE IN DIFFERENT YEASTS USING SUITABLE PROMOTERS.

55

Expression of the α -amylase and glucoamylase gene and secretion of the gene products can be detected in *K. lactis* under control of the promoter of the α -galactosidase (LAC4) gene (Breunig et al., 1984). Expression of the α -amylase and glucoamylase gene and secretion of the both proteins can also be detected in *S. pombe* under control of the ADH1 promoter (Russel et al., 1983) or of the GAL1/10 promoters (Johnston and Davis 1984). After fusion with the GAL1/10 promoter (Johnston and Davis 1984), DHAS promoter (Janowicz et al., 1985), MOX promoter (Leudeboer et al., 1985) or FMDH promoter (EP 87 110 417.0) expression of the α -amylase and glucoamylase

gene and secretion of both proteins can be detected in Hansenula polymorpha.

5 EXPRESSION AND SECRETION STUDIES IN S. OCCIDENTALIS USING THE CELLULASE ENDOGLUCANASE D
(EGD) OF CLOSTRIDIUM THERMOCELLUM AS A MODEL FOR FOREIGN GENE EXPRESSION.

10 As a first example for the expression of a foreign gene in S. occidentalis, the celD gene from C. thermocellum (Millet et al., 1985, Joliff et al., 1986b), which codes for a thermostable cellulase, namely an endoglucanase (EGD) was used. Advantages in using this system are :

15 10 - specific reaction of EGD, which can easily be monitored
- a rapid qualitative assay for colonies with EGD activity (see material and methods)
- availability of antibodies against EGD for specific detection of translation products.

CHARACTERISATION OF THE GLUCOAMYLASE PROMOTER AFTER FUSION WITH THE celD GENE

15 15 Plasmids were constructed containing a replicon from S. occidentalis (SwARS1) (EP 87110370.1) the TRP5 gene from S. cerevisiae (Zalkin and Yanowski, 1982) as a selective marker and, in addition, a GAM/celD gene fusion under the control of different promoters. In the first step the 4.0 kb EcoRI-PvuII fragment from plasmid pBRSwARSGAM (Fig. 6a) was isolated and inserted into the 2296 bp pBR322 EcoRI-PvuII fragment, containing the amp gene and the bacterial origin, resulting in plasmid pBRPGAM (Fig.6a). In addition to the pBR322 sequence this plasmid carries 3.6 kb from the 5' non coding region of the GAM gene and the first 208 bp coding for the N-terminal part (including the signal sequence) of the glucoamylase. A 3.4 kb PvuII fragment from plasmid pCT6O3 (Joliff et al., 1986) containing the coding region of the celD gene missing the 5' 111 bp was inserted into the PvuII single site of pBRPGAM resulting in pBRGC1 (Fig.6b). E. coli transformed with this plasmid results in weak Carboxymethylcellulase (CMCase) activity, analysed by congo red staining (see material and methods) indicating that the glucoamylase promoter is slightly expressed in E. coli. For the construction of a S. occidentalis expression vector the plasmid pCJD5-1 (EP 87110370.1) was cleaved with BamHI/PstI and ligated with the 5.5 kb BglII-PstI fragment or 6.5 kb BamHI-PstI fragment from pBRGC1, resulting in pMPGC1-1 (Fig.6c) and pMPGC1-2 (Fig.6d), respectively. Both plasmids share an in frame fusion of GAM/celD but differ in the length of the 5' upstream region of the GAM gene. S. occidentalis strain NGA58 (trp5 ade) isolated after UV mutagenesis of strain NGA23 was transformed with these plasmids and transformants selected for tryptophan prototrophy and analyzed by means of congo red assay for cellulase activity. Transformants produced active EGD when grown under inducing conditions in 2% maltose (Fig 7a). By analysis with the congo red assay no activity could be detected under repressed conditions in 4% glucose (Fig. 7b). NGA58:pMPGC1-1 and NGA58:pMPGC1-2 transformants secrete active EGD into the culture supernatant when grown in maltose under induced conditions. The intact or complete glucoamylase promoter resides on plasmid pMPGC1-2 as shown by 20 times higher activity compared to the shorter promoter fragment in gene fusion of pMPGC1-1 (table 1). However the 0.3 kb promoter fragment still harbours regulatory units which lead to catabolite repression by glucose (Fig.7b). Comparison of the N-terminal GAM sequence to the signal peptidase cleavage sites of several other secretory proteins (Von Heijne 1983) reveals three potential signal peptidase cleavage sites after the amino acid residues Ala 19, Ala 25 and Ala 34. For determination of the N-terminal amino acid of the mature glucoamylase the protein was isolated for N-terminal amino acid sequencing. However the N-terminus is blocked and therefore could not be sequenced. To confirm the correct cleavage site of the GAM signal sequence in S. occidentalis, the regions between the potential signal peptidase cleavage sites and the PvuII gene fusion site are deleted in the GAM/celD fusion by in vitro mutagenesis to study the influence on secretion of active EGD.

45 45 STUDIES OF THE USE OF SUITABLE S. CEREVISIAE PROMOTERS IN THE EXPRESSION OF FOREIGN GENES IN S. OCCIDENTALIS.

50 The ADH1 promoter (Hitzeman et al., 1981) was isolated as a 1.45 kb BamHI-Sall fragment from plasmid pFM2-1 (Müller et al., 1987) and ligated into plasmid pCJD5-1 (EP 87 110370.1) after cleavage with BamHI and Sall resulting in pMPADH1 (Fig.8a). The GAM gene was excised as a Sall fragment from pJDcg-15 (see below) and inserted into the single Sall site of pMPADH1 (Fig. 8b) resulting in plasmid pMPAG (Fig. 8b). By the means of in vivo recombination (Ma et al., 1987) the GAM promoter from pMPGC1-2 (Fig 6d.) was replaced by the ADH1 promoter of pMPAG leading to plasmid pMPAGC (Fig. 8c). Plasmid pMPGC1-2 was linearized with BamHI; the unique BamHI site of pMPGC1-2 is localized between the TRP5 coding sequence and the glucoamylase promoter. The 2.75 kb Clal fragment of pMPAG containing the ADH1 promoter shares homologies to the TRP5 gene as well as to the GAM coding region up to position +208. For in vivo recombination NGA58 was transformed with BamHI linearized pMPGC1-2 together with approx. 10 fold excess of Clal fragment using the transformation protocol of Ito et al., 1983.

Recombinants were screened for their ability to hydrolyze CMC on 4% glucose. Under these conditions celD expression is repressed in transformants harbouring recircularized pMPGC1-2. From 146 transformants 7 colonies expressing cellulase in the presence of 4% glucose were isolated. NGA58 containing the "in vivo" constructed plasmid pMPAGC expresses active EGD constitutively (see Fig. 7). This result demonstrates that the newly constructed plasmid pMPADH1 serves as a vector for foreign gene expression.

In another experiment a recombinant plasmid was constructed by *in vivo* recombination to study the expression of the GAM/celD gene fusion under control of the PDC promoter (Das and Hollenberg, 1982). Therefore the 6.0 kb SphI-Aval fragment from plasmid pJDcg-15 (EP 87 110370.1) carrying the glucoamylase gene fused to the PDC promoter was subcloned into the respective restriction sites of YRpJD2 (EP 87 110370.1) resulting in plasmid pMPPG (Fig. 9a).

For *in vivo* recombination pMPPG was cleaved with *Cl* and used together with *Bam*HI linearized pMPGC1-2 for transformation of NGA58. Homologous recombination results in plasmid pMPPGC (Fig. 9b) harbouring a GAM-celD fusion under control of the PDC promoter. This construction leads to an expression in *Schwanniomyces occidentalis* NGA58 inducible by glucose and partly repressible by maltose (Table 1). Transformants were analysed for expression of active EGD using the qualitative congo red assay (Fig. 7). The results indicate that the 0.3 kb GAM promoter fragment leads to regulated expression, however less efficient than the 1.3 kb fragment (for comparison see table 1 and Fig. 2A). Expression under the ADH1 promoter leads to efficient constitutive expression. The expression level is comparable with that of induced expression directed by the 1.3 kb GAM promoter.

The fact that plasmids can be constructed by *in vivo* recombination is a proof for homologous recombination and hence offers the possibility for stable site directed integration into the genome of *S. occidentalis*, a prerequisite for the industrial production of foreign proteins.

The expression of the α -amylase gene and glucoamylase gene and secretion of the both gene products in *S. occidentalis* could also be achieved under control of the phosphoglycerate kinase (PGK) promoter (Dobson et al., 1982), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter (Holland and Holland 1979), copperchelatin (CUP1) promoter (Karin et al., 1984, Butt et al., 1984) and alcoholdehydrogenase 2 (ADR2 or ADH2) promoter (Russell et al., 1983, Beier et al., 1985) or galactokinase (GAL1/10) promoter (Johnston and Davis, 1984).

Construction of pJDcg-15 :

The GAM gene was isolated as a 3.2 kb *Bam*HI-Sall DNA fragment from plasmid pMaGAM B/S (Fig. 4) and inserted into the respective restriction sites of plasmid pBM272 Johnston & Davis, 1984). The PDC promoter was isolated from plasmid pCP202 (Kellermann & Hollenberg, 1988) as a Sall-SphI fragment and ligated into the respective single sites of vector YRp7 (Struhl et al., 1979). To be able to isolate the PDC1 promoter as a *Bam*HI fragment the 3.2 kb Sall of pMaGAM B/S containing the GAM gene was fused to the PDC promoter via the single Sall site. The resulting *Bam*HI fragment containing the PDC-promoter was ligated into the single *Bam*HI site of pBM272 containing the GAM gene. The final plasmid is called pJDcg-15 (Fig. 8b).

EXPRESSION OF HUMAN GRANULOCYTE/MACROPHAGE COLONY STIMULATING FACTOR (hGM-CSF)

The gene coding for hGM-CSF (Cantrell et al., 1985) was obtained from British Biotechnology Ltd. as a HindIII-EcoRI fragment. Position +52 of hGM-CSF gene was fused to position + 102 of the GAM signal sequence under control of the ADH1 or GAM promoter. This fusion results in a hybrid protein in which the endopeptidase cleavage site of the GAM signal sequence is linked with a Met included in the synthetic gene offered by British Biotechnology Ltd., followed by the first amino acid of mature hGM-CSF (Ala 18). This fusion was inserted into a *S. occidentalis* expression vector and transformed into NGA58. Western blot analysis revealed secretion of hGM-CSF protein.

EXPRESSION OF INTERLEUKINE-2

The gene coding for IL-2 (Taniguchi et al., 1983) was obtained from British Biotechnology Ltd., as a EcoRI-HindIII fragment. Position +61 of IL-2 gene was fused to position + 102 of the GAM signal sequence under control of the ADH1 or GAM promoter. This fusion results in a hybrid protein in which the endopeptidase cleavage site of the GAM signal sequence is linked with a Met included in the synthetic gene offered by British Biotechnology Ltd., followed by the first amino acid of mature IL-2 (Ala 21). This fusion was inserted into a *S. occidentalis* expression vector and transformed into NGA58. Western blot analysis revealed secretion of IL-2 protein.

ISOLATION OF AN ADDITIONAL SwARS ELEMENT AND A NEW SELECTABLE MARKER

A new SwARS sequence (SwARS2), localized downstream of the coding region of the GAM gene, could be iden-

tified. To analyze the function of SwARS2 a vector was constructed containing the Trps gene and SwARS2 in pBR322, resulting in pMPTS2-A and pMPTS2-B, respectively (Fig. 10). For the construction of pMPTS2-A and pMPTS2-B the 5.5 kb BamHI - BglI fragment of pBRSwARSGAM, constructed by insertion of the EcoRI-fragment shown in Fig.2a into the single EcoRI site of plasmid pBRSwARS1, (disclosed in EP 87 110370.1), was isolated and ligated with 3.2 kb BamHI TRP5 fragment obtained from YRpJD2. High frequency transformation of *S. occidentalis* is obtained using this 5 plasmid.

As an additional selectable marker the LEU2 gene of *S. occidentalis* was isolated. The LEU2 gene, coding for isopropyl malate dehydrogenase, was cloned by functional complementation of the mutant AH22 (his4leu2) (Hinnen et al., 1978) of *S. cerevisiae* by transformation with a cosmid genomic library from *S. occidentalis*.

10 The LEU2 gene was found to reside on a 9 kb EcoRI fragment of *Schwanniomyces* genomic DNA; its function could be abolished by insertion of the TRP5 gene of *S. cerevisiae*. The disrupted LEU2 gene was integrated into the genome of NGA58 by homologous recombination resulting in a new strain NGA581 (TRP5, LEU2, and ADE). An analogous experiment with the cloned HIS4 gene of *S. occidentalis* (EP 87 110 370.1) led to a strain NGA58h (TRP5, HIS4, ADE).

15 IMPROVED EXPRESSION OF HETEROLOGOUS GENES USING THE T7 RNA POLYMERASE AND T7 PROMOTER AND TERMINATOR FOR THE EXPRESSION OF FOREIGN GENES IN *S. occidentalis*

20 T7 RNA polymerase (Fuerst et al., 1986) can be expressed in *S. occidentalis* after fusion of the structural gene (Dunn and Studier, 1984) using the BamHI restriction site near the translation initiation codon with suitable promoters as α -amylase, GAM (0.3 kb fragment) or PDC and homologous integration into the genome via any homologous cloned 25 gene, e.g. HIS4, LEU2, AMY1 and GAM. A suitable fragment, containing the fusion flanked by homologous cloned sequences can be isolated and used together with pCJD5-1 (Fig. 6c) for co-transformation of NGA58, selecting on YNB containing 0.5% casaminoacids. Colonies containing the integrated T7 RNA polymerase gene can be identified by Southern analysis after isolation of transformants mutated in the target gene.

25 For the expression of foreign genes a *S. occidentalis* expression vector was constructed containing SwARS2, the ColE replication origin of *E. coli*, a suitable selective marker for *S. occidentalis* (e.g. TRPS, HIS4, LEU2) and a cassette containing the T7 polymerase promoter and terminator sequences isolated from plasmid pAR2529 (Fuerst et al., 1986), separated by the polylinker sequence from M13mp19. To study the efficiency of the T7 system we expressed a GAM 30 celD fusion (GAM signal sequence and structural celD gene obtained as a BamHI-Asp 718 fragment from pMPGC1-2) under control of the T7 promoter by insertion into the BamHI-Asp718 site of the polylinker sequence. Using this system active cellulase is secreted into the culture medium.

MATERIALS AND METHODS

35 1. The microorganisms used are as follows :

<i>Schwanniomyces occidentalis</i>	NGA23 : (DSM 3792)
40 <i>Schwanniomyces occidentalis</i>	NGA58
<i>Kluyveromyces lactis</i>	SD11 : (DSM 3795)
45 <i>Schizosaccharomyces pombe</i>	LEU1-32, HIS5-303 (DSM 3796)
<i>Saccharomyces cerevisiae</i>	YNN27 (URA3-52, TRP1-298, <u>GAL2</u>)

50 The *Escherichia coli* strain used is as follows *E. coli* HB101 F-, hsd S20 ($r_B^- m_B^-$), recA13, ara-14, proA2, lacY1, galK2, rpsL20 (Sm^r), xyl-5, mtl-1, supE44)

55 2. Culture media :

Schwanniomyces occidentalis strain NGA58 (isolated by UV mutagenesis of strain NGA23) was grown at 30°C in 55 YNB (0.67% yeast nitrogen base without amino acids) supplemented with appropriate amino acids and with 2% of soluble starch when indicated.

55 *S. cerevisiae* strain YNN27 (URA3-52, TRP1-298, GAL2) (Stinchomb et al 1980); *K. lactis* strain SD11 TRP1 (Das and Hollenberg, 1982); *S. pombe* strains LEU1-32, HIS5-303 (obtained from W.D. Heyer). These strains were grown under conditions as described above for *Schwanniomyces occidentalis* with the replacement of 2% soluble starch by

2% glucose, where indicated.

The culture media for *S. occidentalis* and *Hansenula polymorpha* were buffered with 0.2 M NaP0₄ buffer pH 6.2; those for *S. cerevisiae* and *K. lactis* were buffered with 0.1 M citrate buffer pH 6 and those for *S. pombe* with 0.05 M acetate buffer pH 6.

5 The pH optimum for the cellulase (Joliff et al., 1986a), α -amylase and glucoamylase (Wilson and Ingledew, 1982) is approx. pH 6.

E. coli HB101 (Bolivar et al., 1977) was grown in LB medium (Maniatis et al., 1982) with penicillin G (150 μ g/ml).

3. Miscellaneous methods

10 Plasmid DNA was purified either by CsCl gradient centrifugation (Maniatis et al., 1982) or by rapid alkaline extraction of plasmid DNA described by Birnboim and Doly (1979).

Yeast crude extracts were prepared by the method of Schatz (1979). Yeast minilysates were prepared by the method of Sherman et al., (1983).

15 DNA fragments were isolated by the "low melting agarose" procedure as described by Gafner et al., (1983). Southern hybridization was carried out as described by Southern (1975).

Yeast transformation was performed according to Klebe et al. (1983) or Ito et al. (1983). The transformants or integrants were tested for expression and secretion of cellulase using either the qualitative congo red assay (Teather and Wood, 1982: halo formation on carboxymethylcellulose agar plates) or a quantitative cellulase assay according to 20 Joliff et al. (1986a). One unit (U) is defined as the amount of enzyme that released 1 umol para-nitrophenol per min at 60°C. Western analysis was performed as described by Towbin et al. (1979).

The transformants or integrants were tested for expression and secretion of α -amylase using either a standardized α -amylase enzyme test from Merck, W. Germany or a qualitative starch degradation test (halo formation on starch agar plates after staining with iodine), as follows: in this test the rate of formation of 2-chloro-4-nitrophenol is determined by photometry at 405 nm in 0.1 M potassium phosphate at 37°C. The enzyme unit (U) is defined as the amount of enzyme catalysing the formation of 1 μ mol 2-chloro-4-nitrophenol per min at 37°C.

Glucoamylase activity was measured by a stop assay method:

After incubation of the sample in 10% soluble starch in 0.05M KH₂PO₄-NaOH (pH 5.0) at 50°C the amount of glucose produced is determined by the glucose dehydrogenase method (system glucose Merck). The quantity of NADH formed is proportional to the glucose concentration. The enzyme unit (U) is defined as the amount of enzyme catalysing the formation of 1 umol glucose/min at 50°C.

TABLE 1

	Transformant	OD _{600 nm}	EDG-activity mU/ml
35	NGA58:	2.3	0.12
	pMPGC1-1 (0.3 kb promoter)	5.3	0.73
40	NGA58:	2.0	0.53
	pMPGC1-2 (1.3 kb GAM promoter)	7.9	1.83
45	NGA58:	3.0	0.52
	pMPAGC ADH1 promoter	5.9	1.07
	NGA58:	2.2*	0.40*
	pMPPGC	4.5*	0.64*
	PDC1-promoter	2.9	0.19
		7.2	0.31

Transformants were grown in YNB, 0.5% Casaminoacids, 2% maltose, 20 mg/adenine and 0.2 M NaP buffer pH 6.2

*: values for transformants grown in 4% glucose

50 LIST OF REFERENCES

Beier et al. (1985), Mol. Cell. Biol. 5, 1743-1749.

55 Birnboim & Doly (1979), Nuc. Acids Res. 7, 1513-1523.

Bitter et al. (1984), Proc. Natl. Acad. Sci. USA 81, 5330-5334.

Bolivar et al. (1977), *Gene* 2, 75-93.

Breunig (1984), *NUCl. Acids Res.* 12, 2327-2341.

5 Butt et al. (1984), *Proc. Natl. Acad. Sci. USA* 81, 3332-3336.

Cantrell et al. (1985), *Proc. Natl. Acad. Sci USA* 82, 6250-6254.

10 Das & Hollenberg (1982), *Curr. Genet.* 6, 123-128.

Dobson et al. (1982), *Nucleic Acids Research* 10, 2625-2637.

Dunn & Studier (1984), *Proc. Natl. Acad. Sci. USA* 81, 2035-2039.

15 Emr et al. (1983), *Proc. Natl. Acad. Sci. USA* 80, 7080-7084.

Fuerst et al. (1986), *Proc. Natl. Acad. Sci. USA* 83, 8122-8126.

20 Gafner et al. (1983), *Embo J.* 2, 583-591.

Hinnen et al. (1978), *Proc. Natl. Acad. Sci. USA* 75, 1929-1933.

Hitzeman et al. (1981), *Nature* 294, 717-722.

25 Holland & Holland (1979), *J. Biol. Chem.* 254, 5466-5474.

Ito et al. (1983) *J. Bacteriol.* 153, 163-168.

30 Janowicz et al. (1985), *Nucleic Acids Res.* 9, 3043-3062.

Johnston & Davis (1984), *Mol. Cell. Biol.* 4, 1440-1448.

Joliff et al. (1986a), *Bio/Technology* 4, 896-900.

35 Joliff et al. (1986b), *NUCl. Acids Res.* 14, 8605-8613.

Karin et al. (1984), *Proc. Natl. Acad. Sci. USA* 81, 337-341.

Kellermann & Hollenberg (1988), *Curr. Genetics* 14, 337-344.

40 Klebe et al. (1983), *Gene* 25, 333-341.

Kramer et al. (1984), *Nucl. Acids. Res.* 12, 9441-9456.

45 Kurjan & Herskowitz (1982), *Cell* 30, 933-943.

Ledeboer et al. (1985), *Nucl. Acids. Res.* 9, 3063-3082.

50 Ma et al. (1987), *Gene* 8, 201-216.

Maniatis, Fritsch & Sambrook (1982), *Molecular cloning: A laboratory manual*, Cold Spring Harbour Laboratory press, New York.

Millet et al. (1985), *FEMS Microbiol. Lett.* 29, 145-149.

55 Müller et al. (1987), *Mol. Gen. Genetics* 207, 421-429.

Oteng-Gyang et al. (1981), *Zeitschrift allgemeine Mikrobiologie* 21, 537-544.

Perlman et al. (1983), J. Mol. 167, 391-409.

Price et al. (1978), Microbiol. Rev. 42, 161-193.

5 Russell et al. (1983), J. Biol. Chem. 258, 2674-2682.

Sanger et al. (1977), Proc. Natl. Acad. Sci. USA 74, 5 463-5 467.

10 Schatz (1979), Meth. Enzymol. LVI, 40-50.

Sharp et al. (1980), in: Methods Enzymol. 65,

Sherman et al. (1983), Methods in yeast genetics in C.S.H. Laboratory, C.S.H. New York, 119.

15 Sills & Stewart (1982), J. Inst. Brew. 88, 313-316.

Sills et al. (1984 a), J. Inst. Brew. 90, 311-314.

Sills et al. (1984 b), APPL. Microb. Biotechnol. 20, 124-128.

20 Southern (1975), J. Mol. Biol. 98, 503-517.

Stinchcomb et al. (1980), Proc. Natl. Acad. Sci. USA 77, 4 559-4 563.

25 Struhl et al. (1979), Proc. Natl. Acad. Sci. USA 76, 1035-1039.

Taniguchi et al. (1983), Nature 302, 305-309.

Teather & Wood (1982), Appl. Env. Microbiol. 43, 777-780.

30 Towbin et al. (1979), Proc. Natl. Acad. Sci. USA 76, 4350-4354.

von Heijne (1983), Eur. J. Biochem. 133, 17-21.

35 Wilson & Ingledew (1982), Appl. Environ. Microbiol. 44, 301-307.

Zalkin & Yanowsky (1982), J. Biol. Chem. 257, 1491-1500.

40 **Claims**

Claims for the following Contracting States : AT, BE, CH, LI, DE, FR, GB, GR, IT, LU, NL, SE

45 1. A yeast cell of the genus Schwanniomyces,
characterized in that
said yeast cell contains at least one expression cassette comprising :

50 a) a first DNA sequence serving as a regulon,
b) optionally a second DNA sequence coding for a signal peptide
55 c) a third DNA sequence coding for a foreign protein, and
d) optionally a fourth DNA sequence serving as a terminator,
wherein the first, second and fourth DNA sequence are derived from yeast.

2. The yeast cell according to claim 1,
characterized in that
 said first and/or second and/or fourth DNA sequence are derived from a gene coding for an amylolytic enzyme.

5 3. The yeast cell according to claim 2,
characterized in that
 said amylolytic enzyme is an α -amylase or glucoamylase.

10 4. The yeast cell according to any of claims 1 to 3,
characterized in that
 said first and/or second and/or fourth DNA sequences are derived from a yeast of a member of the genera Debar-
iomycetes, Saccharomycopse, Lipomyces, Pichia, or Saccharomyces, preferably Schwanniomyces.

15 5. The yeast cell according to claim 4,
characterized in that
 said yeast is Schwanniomyces occidentalis.

6. The yeast cell according to any of claims 1 to 5,
characterized in that
 said first DNA sequence is comprised by a 1.8 kb BglII-Xhol-fragment, if derived from Schwanniomyces occidentalis
 α -amylase gene, or a 1.3 kb BamHI-PvuII fragment if derived from Schwanniomyces occidentalis glucoamylase
 gene.

25 7. The yeast cell according to claim 5,
characterized in that
 said first DNA sequence is comprised by part or all of the DNA sequence corresponding to bases -1 to -540 of the
 region preceding the structural gene coding for α -amylase or part or all of the DNA sequence corresponding to
 bases -1 to -320 of the region preceding the structural gene coding for glucoamylase.

30 8. The yeast cell according to any of claims 1 to 5,
characterized in that
 said first DNA sequence corresponds to one of the regulons regulating expression of:
 a) ADH1,ADH2, PDC1, GAL1/10, PGK, GAPDH, wherein the regulon is preferably obtained from Saccharo-
myces cerevisiae,
 b) LAC4, wherein the regulon is preferably obtained from Kluyveromyces lactis, or
 c) corresponding regulons, preferably obtained from Schwanniomyces.

40 9. The yeast cell according to any of claims 1 to 8, **characterized in that** said second DNA sequence is coding for
 all or part of at least one of the following peptides :

45 a) **Met Arg Phe Ser Thr Glu Gly Phe Thr Ser Lys Val Val Ala Ala Ile**
Leu Ala Phe Ser Arg Leu Val Ser Ala;

50 b) **Met Arg Phe Ser Thr Glu Gly Phe Thr Ser Lys Val Val Ala Ala Ile**
Leu Ala Phe Ser Arg Leu Val Ser Ala Gln Pro Ile Ile Phe Asp Met Arg;

c) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu
Val Ser Ala;

5

d) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu
Val Ser Ala Ile Gln Ala Ala Pro Ala;

10

e) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu
Val Ser Ala Ile Gln Ala Ala Pro Ala Ser Ser Ile Gly Ser Ser Ala Ser Ala

15

10. A yeast cell according to claim 9,
characterized in that
20 the DNA sequence coding for said peptide corresponds to part or all of one of the following DNA sequences:

25 a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT

TTAGCATTCTCAAGATTGGTATCTGCT,

25

30 b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TAGCATTCTCAAGATTGGTATCTGCTCAACCGATTATTTTGACATGAGA;

35

c) ATGATTTCTGAAGCTGATTAAAGTATAGTAATTGGTTGGATT
GTTAGTGCT;

40

d) ATGATTTCTGAAGCTGATTAAAGTATAGTAATTGGTTGGATT
GTTAGTGCTATCCAAGCAGCCCCTGCC;

45

e) ATGATTTCTGAAGCTGATTAAAGTATAGTAATTGGTTGGATT
GTTAGTGCTATCCAAGCAGCCCCTGCCCTTCGATTGGATCTAGTGCTTCA
GCA.

50 11. The yeast cell according to any of claims 1 to 10,

characterized in that

said third DNA sequence coding for a foreign protein is a natural or synthetic DNA sequence encoding a cellulase, interleukin, insulin like growth factor, interferon, lymphokine, human growth factor, nerve growth factor, aprotinin, insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines or human granulocyte/macrophage colony stimulating factor.

55 12. The yeast cell according to any of claims 1 to 11,

characterized in that

said fourth DNA sequence originates from a yeast of a member of the genera Saccharomyces, Pichia, Hansenula, preferably Schwanniomyces.

13. The yeast cell according to claim 12,
 5 **characterized in that**
 said fourth DNA sequence corresponds to part or all of the terminator comprised by nucleotides 1537 to 1740 of the Schwanniomyces α -amylase gene or part or all of the terminator comprised by nucleotides 2875 to 3320 of the Schwanniomyces glucoamylase gene.

10 14. The yeast cell according to any of claims 1 to 13,
characterized in that
 said first and/or second and/or third and/or fourth DNA sequence has been modified by insertion or deletion or substitution of one or more nucleotides while the biological activity of the DNA sequence is retained or improved.

15 15. The yeast cell according to any of claims 1 to 14,
characterized in that
 said expression cassette is carried by a circular or linear vector, which optionally comprises one or more selective marker genes.

20 16. The yeast cell according to claim 15,
characterized in that
 said vector comprises any of the following selective marker genes : TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, preferably obtained from Saccharomyces, Hansenula Pichia or Schwanniomyces, AMY1 or GAM1, preferably obtained from Schwanniomyces.

25 17. The yeast cell according to any of claims 15 and 16,
characterized in that
 said vector further comprises a DNA sequence capable of controlling autonomous replication and stable maintenance of the vector in the host organism.

30 18. The yeast cell according to claim 17
characterized in that
 said DNA sequence is derived from the genome of a member of the genera Saccharomyces, Pichia, Hansenula or Kluyveromyces, preferably Schwanniomyces.

35 19. A yeast cell according to any of claims 17 or 18,
characterized in that
 said DNA sequence is selected from the autonomously replicating sequences ARS, preferably an ARS derived from the respective host organism.

40 20. The yeast cell according to claim 19,
characterized in that
 said DNA sequence is SwARS1 or SwARS2, capable of controlling autonomous replication and stable maintenance of the vector in Schwanniomyces, wherein SwARS1 is comprised on a *Clal*-*Bam*HI fragment adjacent to the *Eco*RI fragment shown in Fig. 2a, and SwARS2 is comprised on an *Eco*RI-*Bgl*III fragment contained in the *Eco*RI fragment shown in Fig. 2a.

45 21. The yeast cell according to any of claims 17 to 20,
characterized in that
 said DNA sequence has been modified by deletion, insertion or substitution of one or more nucleotides, while retaining or improving its biological function.

50 22. The yeast cell according to any of claims 1 to 21,
characterized in that
 said expression cassette is integrated into the genome, optionally in multiple copies, preferably via homologous recombination.

55 23. A process for the production of polypeptide, wherein a yeast host organism is cultivated under suitable conditions

and the polypeptide is recovered in a manner known per se,
characterized in that
 said yeast host organism is a yeast cell according to any of claims 1 to 22.

5 24. The process according to claim 23
characterized in that
 yeast cells of the species Schwanniomyces occidentalis are used as a host organism and are preferably cultivated on starch dextrin, maltose and/or plant biomass.

10 25. The process according to any of claims 23 or 24,
characterized in that
 yeast cells of the genus Schwanniomyces are used for the production of single cell protein.

15 26. The process according to any of claims 23 to 25,
characterized in that
 said polypeptide is amylase or glucoamylase.

27. Use of a yeast cell according to any of claim 1 to 22 for transformation and expression of foreign genes.

20 **Claims for the following Contracting State: ES**

1. A process for preparing a yeast cell capable of expressing foreign protein,
characterized in that
 25 at least one expression cassette comprising:

- a first DNA sequence serving as a regulon,
- 30 b) optionally a second DNA sequence coding for a signal peptide
- c) a third DNA sequence coding for a foreign protein, and
- d) optionally a fourth DNA sequence serving as a terminator,

 35 is inserted into a yeast cell of the genus Schwanniomyces, wherein the first, second and fourth DNA sequence are derived from yeast.

2. The process according to claim 1,
characterized in that
 40 said first and/or second and/or fourth DNA sequence are derived from a gene coding for a amylolytic enzyme.

3. The process according to claim 2,
characterized in that
 45 said amylolytic enzyme is an α -amylase or glucoamylase.

4. The process according to any of claims 1 to 3,
characterized in that
 50 said first and/or second and/or fourth DNA sequences are derived from a yeast of a member of the genera Debaromyces, Saccharomyces, Lipomyces, Pichia, or Saccharomyces, preferably Schwanniomyces.

5. The process according to claim 4,
characterized in that
 55 said yeast is Schwanniomyces occidentalis.

6. The process according to any of claims 1 to 5,
characterized in that
 said first DNA sequence is comprised by a 1.8 kb BglII-Xhol-fragment, if derived from Schwanniomyces occidentalis α -amylase gene, or a 1.3 kb BamHI-PvuII fragment if derived from Schwanniomyces occidentalis glucoamylase

gene.

7. The process according to claim 5,
characterized in that

5 said first DNA sequence is comprised by part or all of the DNA sequence corresponding to bases -1 to -540 of the region preceding the structural gene coding for α -amylase or part or all of the DNA sequence corresponding to bases -1 to -320 of the region preceding the structural gene coding for glucoamylase.

8. The process according to any of claims 1 to 5,

10 **characterized in that**
 said first DNA sequence corresponds to one of the regulons regulating expression of:

15 a) ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, wherein the regulon is preferably obtained from Saccharomyces cerevisiae,

b) LAC4, wherein the regulon is preferably obtained from Kluyveromyces lactis, or

c) corresponding regulons, preferably obtained from Schwanniomyces.

20 9. The process according to any of claims 1 to 8,

characterized in that

said second DNA sequence is coding for all or part of at least one of the following peptides:

25 a) **MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle**
LeuAlaPheSerArgLeuValSerAla;

30 b) **MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle**
LeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;

35 c) **MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu**
ValSerAla;

40 d) **MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu**
ValSerAlaIleGlnAlaAlaProAla;

45 e) **MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu**
ValSerAlaIleGlnAlaAlaProAlaSerSerIleGlySerSerAlaSerAla

50 10. A process according to claim 9,

characterized in that

the DNA sequence coding for said peptide corresponds to part or all of one of the following DNA sequences:

55 a) **ATGAGATTTCACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT**
TTAGCATTCTCAAGATTGGTATCTGCT,

5
b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TAGCATTCTCAAGATTGGTATCTGCTAACCGATTATTTGACATGAGA;

5

10
c) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGATTA
GTTAGTGCT;

10

15
d) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGATTA
GTTAGTGCTATCCAAGCAGCCCTGCC;

15

20
e) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGATTA
GTTAGTGCTATCCAAGCAGCCCTGCCCTTCGATTGGATCTAGTGCTTC
GCA.

25 11. The process according to any of claims 1 to 10,
characterized in that
said third DNA sequence coding for a foreign protein is a natural or synthetic DNA sequence encoding a cellulase, interleukin, insulin like growth factor, interferon, lymphokine, human growth factor, nerve growth factor, aprotinin, insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines or human granulocyte/macrophage colony stimulating factor.

30 12. The process according to any of claims 1 to 11,
characterized in that
said fourth DNA sequence originates from a yeast of a member of the genera Saccharomyces, Pichia, Hansenula, preferably Schwanniomyces.

35 13. The process according to claim 12,
characterized in that
said fourth DNA sequence corresponds to part or all of the terminator comprised by nucleotides 1537 to 1740 of the Schwanniomyces α -amylase gene or part or all of the terminator comprised by nucleotides 2875 to 3320 of the Schwanniomyces glucoamylase gene.

40 14. The process according to any of claims 1 to 13,
characterized in that
said first and/or second and/or third and/or fourth DNA sequence has been modified by insertion or deletion or substitution of one or more nucleotides while the biological activity of the DNA sequence is retained or improved.

45 15. The process according to any of claims 1 to 14,
characterized in that
said expression cassette is carried by a circular or linear vector, which optionally comprises one or more selective marker genes.

50 16. The process according to claim 15,
characterized in that
said vector comprises any of the following selective marker genes : TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, preferably obtained from Saccharomyces, Hansenula Pichia or Schwanniomyces, AMY1 or GAM1, preferably obtained from Schwanniomyces.

55 17. The process according to any of claims 15 and 16,

characterized in that

said vector further comprises a DNA sequence capable of controlling autonomous replication and stable maintenance of the vector in the host organism.

5 18. The process according to claim 17

characterized in that

said DNA sequence is derived from the genome of a member of the genera Saccharomyces, Pichia, Hansenula or Kluyveromyces, preferably Schwanniomyces.

10 19. A process according to any of claims 17 or 18,

characterized in that

said DNA sequence is selected from the autonomously replicating sequences ARS, preferably an ARS derived from the respective host organism.

15 20. The process according to claim 19,

characterized in that

said DNA sequence is SwARS1 or SwARS2, capable of controlling autonomous replication and stable maintenance of the vector in Schwanniomyces, wherein SwARS1 is comprised on a ClaI-BamHI fragment adjacent to the EcoRI fragment shown in Fig. 2a, and SwARS2 is comprised on an EcoRI-BglII fragment contained in the EcoRI fragment shown in Fig 2a.

20

21. The process according to any of claims 17 to 20,

characterized in that

said DNA sequence has been modified by deletion, insertion or substitution of one or more nucleotides, while retaining or improving its biological function.

25

22. The process according to any of claims 1 to 21,

characterized in that

said expression cassette is integrated into the genome, optionally in multiple copies, preferably via homologous recombination.

30

23. A process for the production of polypeptide, wherein a yeast host organism is cultivated under suitable conditions and the polypeptide is recovered in a manner known per se,

characterized in that

said yeast host organism is a yeast cell according to any of claims 1 to 22.

35

24. The process according to claim 23

characterized in that

yeast cells of the species Schwanniomyces occidentalis are used as a host organism and are preferably cultivated on starch dextrin, maltose and/or plant biomass.

40

25. The process according to any of claims 23 or 24,

characterized in that

yeast cells of the genus Schwanniomyces are used for the production of single cell protein.

45

26. The process according to any of claims 23 to 25,

characterized in that

said polypeptide is amylase or glucoamylase.

50

27. Use of a yeast cell according to any of claim 1 to 22 for transformation and expression of foreign genes.

Patentansprüche

55

Patentansprüche für folgende Vertragsstaaten: AT, BE, CH, LI, DE, FR, GB, GR, IT, LU, NL, SE

1. Hefezelle der Gattung Schwanniomyces, dadurch gekennzeichnet, daß die Hefezelle mindestens eine Expressi-

onskassette enthält, umfassend:

- a) eine erste DNA-Sequenz, die als Regulon dient,
- 5 b) gegebenenfalls eine zweite DNA-Sequenz, die für ein Signalpeptid kodiert,
- c) eine dritte DNA-Sequenz, die für ein Fremdprotein kodiert, und
- 10 d) gegebenenfalls eine vierte DNA-Sequenz, die als ein Terminator dient,

wobei die erste, zweite und vierte DNA-Sequenz von Hefe abgeleitet sind.

- 15 2. Hefezelle nach Anspruch 1, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder vierte DNA-Sequenz von einem für ein amylolytisches Enzym kodierenden Gen abgeleitet sind.
3. Hefezelle nach Anspruch 2, dadurch gekennzeichnet, daß das amylolytische Enzym eine α -Amylase oder Glucoamylase ist.
- 20 4. Hefezelle nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß die ersten und/oder zweiten und/oder vierten DNA-Sequenzen von einer Hefe eines Angehörigen der Gattungen Debariomyces, Saccharomyces, Lipomyces, Pichia oder Saccharomyces, bevorzugt Schwanniomyces, abgeleitet sind.
5. Hefezelle nach Anspruch 4, dadurch gekennzeichnet, daß die Hefe Schwanniomyces occidentalis ist.
- 25 6. Hefezelle nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz von einem 1,8 kb BglII-Xhol-Fragment umfaßt ist, wenn sie vom Schwanniomyces occidentalis- α -Amylasegen abgeleitet ist, oder von einem 1,3 kb BamHI-PvuII-Fragment, wenn sie von einem Schwanniomyces occidentalis-Glucoamylasegen abgeleitet ist.
- 30 7. Hefezelle nach Anspruch 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz von einem Teil oder der gesamten DNA-Sequenz umfaßt ist, die den Basen -1 bis -540 des Bereiches entspricht, der dem für α -Amylase kodierenden Strukturen vorangeht, oder einem Teil oder der Gesamtheit der DNA-Sequenz, die den Basen -1 bis -320 des Bereiches entspricht, der dem für Glucoamylase kodierenden Strukturen vorangeht.
- 35 8. Hefezelle nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz einem der Regulons entspricht, die die Expression regulieren von
 - a) ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, wobei das Regulon bevorzugt von Saccharomyces cerevisiae erhalten ist,
 - 40 b) LAC4, wobei das Regulon bevorzugt von Kluyveromvces lactis erhalten ist, oder
 - c) entsprechenden Regulons, bevorzugt aus Schwanniomyces erhalten.
- 45 9. Hefezelle nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß die zweite DNA-Sequenz für die Gesamtheit oder einen Teil mindestens eines der folgenden Peptide kodiert:
 - 50 a) **MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAla;**
 - b) **MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;**

c) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu
Val Ser Ala;

5

d) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu
Val Ser Ala Ile Gln Ala Ala Pro Ala;

10

e) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu
Val Ser Ala Ile Gln Ala Ala Pro Ala Ser Ser Ile Gly Ser Ser Ala Ser Ala

15

10. Hefezelle nach Anspruch 9, dadurch gekennzeichnet, daß die für das Peptid kodierende DNA-Sequenz einem Teil
oder der Gesamtheit einer der folgenden DNA-Sequenzen entspricht:

20

a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TTAGCATTCTCAAGATTGGTATCTGCT,

25

b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TAGCATTCTCAAGATTGGTATCTGCTCAACCGATTATTTTGACATGAGA;

30

c) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCT;

35

d) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCC;

40

e) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCCCTTCGATTGGATCTAGTGCTTCA
GCA.

45

11. Hefezelle nach einem der Ansprüche 1 bis 10, dadurch gekennzeichnet, daß die dritte DNA-Sequenz, die für ein
Fremdprotein kodiert, eine natürliche oder synthetische DNA-Sequenz ist, die für eine Zellulase, Interleukin, insu-
linähnlichen Wachstumsfaktor, Interferon, Lymphokin, menschlichen Wachstumsfaktor, Nervenwachstumsfaktor,
Aprotinin, Insulin, Hirudin, Hormone, Blutgerinnungsfaktoren, Hepatitis-B-Oberflächen- oder -Kernantigene, virale
oder bakterielle Impfstoffe oder humanen Granulocyten/Makrophagen-Kolonie-stimulierenden Faktor kodiert.

55

12. Hefezelle nach einem der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß die vierte DNA-Sequenz von einer
Hefe eines Angehörigen der Gattungen Saccharomyces, Pichia, Hansenula, bevorzugt Schwanniomyces, stammt.

13. Hefezelle nach Anspruch 12, dadurch gekennzeichnet, daß die vierte DNA-Sequenz einem Teil oder der Gesamt-

heit des Terminators entspricht, der von den Nukleotiden 1537 bis 1740 des Schwanniomyces- α -Amylasegenes umfaßt ist, oder einem Teil oder der Gesamtheit des Terminators, der von den Nukleotiden 2875 bis 3320 des Schwanniomyces-Glucoamylasegenes umfaßt ist.

5 14. Hefezelle nach einem der Ansprüche 1 bis 13, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder dritte und/oder vierte DNA-Sequenz durch Insertion oder Deletion oder Substitution eines oder mehrerer Nukleotide modifiziert worden ist, während die biologische Aktivität der DNA-Sequenz beibehalten oder verbessert wird.

10 15. Hefezelle nach einem der Ansprüche 1 bis 14, dadurch gekennzeichnet, daß die Expressionskassette von einem zirkulären oder linearen Vektor getragen wird, der gegebenenfalls ein oder mehrere selektive Markergene umfaßt.

15 16. Hefezelle nach Anspruch 15, dadurch gekennzeichnet, daß der Vektor eines der folgenden selektiven Markergene umfaßt:
TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, bevorzugt erhalten von Saccharomyces, Hansenula, Pichia oder Schwanniomyces, AMY1 oder GAM1, bevorzugt erhalten von Schwanniomyces.

20 17. Hefezelle nach einem der Ansprüche 15 und 16, dadurch gekennzeichnet, daß der Vektor weiter eine DNA-Sequenz umfaßt, die die autonome Replikation und stabile Beibehaltung des Vektors in dem Wirtsorganismus steuern kann.

25 18. Hefezelle nach Anspruch 17, dadurch gekennzeichnet, daß die DNA-Sequenz von dem Genom eines Angehörigen der Gattungen Saccharomyces, Pichia, Hansenula oder Kluyveromyces abgeleitet ist, bevorzugt Schwanniomyces.

30 19. Hefezelle nach einem der Ansprüche 17 oder 18, dadurch gekennzeichnet, daß die DNA-Sequenz aus den autonomen replizierenden Sequenzen ARS, bevorzugt einer ARS, die von dem jeweiligen Wirtsorganismus abgeleitet ist, ausgewählt ist.

35 20. Hefezelle nach Anspruch 19, dadurch gekennzeichnet, daß die DNA-Sequenz SwARS1 oder SwARS2 ist, die die autonome Replikation und stabile Beibehaltung des Vektors in Schwanniomyces steuern können, wobei SwARS1 von einem Clal-BamHI-Fragment umfaßt ist, das dem in Figur 2a gezeigten EcoRI-Fragment benachbart ist, und SwARS2 von einem EcoRI-BglIII-Fragment umfaßt ist, das in dem in Figur 2a gezeigten EcoRI-Fragment enthalten ist.

40 21. Hefezelle nach einem der Ansprüche 17 bis 20, dadurch gekennzeichnet, daß die DNA-Sequenz durch Deletion, Insertion oder Substitution eines oder mehrerer Nukleotide modifiziert worden ist, während ihre biologische Funktion beibehalten oder verbessert ist.

45 22. Hefezelle nach einem der Ansprüche 1 bis 21, dadurch gekennzeichnet, daß die Expressionskassette in das Genom integriert wird, gegebenenfalls in mehreren Kopien, bevorzugt über homologe Rekombination.

50 23. Verfahren zum Erzeugen von Polypeptid, wobei ein Hefewirtsorganismus unter geeigneten Bedingungen kultiviert und das Polypeptid in an sich bekannter Weise gewonnen wird, dadurch gekennzeichnet, daß der Hefezorganismus eine Hefezelle gemäß einem der Ansprüche 1 bis 22 ist.

55 24. Verfahren nach Anspruch 23, dadurch gekennzeichnet, daß Hefezellen der Art Schwanniomyces occidentalis als ein Wirtsorganismus verwendet werden und bevorzugt auf Stärke, Dextrin, Maltose und/oder Pflanzenbiomasse kultiviert werden.

50 25. Verfahren nach einem der Ansprüche 23 oder 24, dadurch gekennzeichnet, daß Hefezellen der Gattung Schwanniomyces für die Erzeugung von Einzelzellprotein verwendet werden.

55 26. Verfahren nach einem der Ansprüche 23 bis 25, dadurch gekennzeichnet, daß das Polypeptid Amylase oder Glycoamylase ist.

55 27. Verwendung einer Hefezelle nach einem der Ansprüche 1 bis 22 zur Transformation und Expression von Fremdgenen.

Patentansprüche für folgenden Vertragsstaat: ES

1. Verfahren zum Herstellen einer Hefezelle, die Fremdprotein exprimieren kann, dadurch gekennzeichnet, daß mindestens eine Expressionskassette in eine Hefezelle der Gattung Schwanniomyces insertiert wird, wobei die Expressionskassette umfaßt:
 - a) eine erste DNA-Sequenz, die als Regulon dient,
 - b) gegebenenfalls eine zweite DNA-Sequenz, die für ein Signalpeptid kodiert,
 - c) eine dritte DNA-Sequenz, die für ein Fremdprotein kodiert, und
 - d) gegebenenfalls eine vierte DNA-Sequenz, die als Terminator dient,
 wobei die erste, zweite und vierte DNA-Sequenz von Hefe abgeleitet sind.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder vierte DNA-Sequenz von einem für ein amylolytisches Enzym kodierenden Gen abgeleitet sind.
3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß das amylolytische Enzym eine α -Amylase oder Glucoamylase ist.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß die ersten und/oder zweiten und/oder vierten DNA-Sequenzen von einer Hefe eines Angehörigen der Gattungen Debariomyces, Saccharomyces, Lipomyces, Pichia oder Saccharomyces, bevorzugt Schwanniomyces, abgeleitet sind.
5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß die Hefe Schwanniomyces occidentalis ist.
6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz von einem 1,8 kb BgIII-Xhol-Fragment umfaßt ist, wenn sie vom Schwanniomyces occidentalis- α -Amylasegen abgeleitet ist, oder von einem 1,3 kb BamHI-PvuII-Fragment umfaßt ist, wenn sie vom Schwanniomyces occidentalis-Glucoamylasegen abgeleitet ist.
7. Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz von einem Teil oder der Gesamtheit der DNA-Sequenz umfaßt ist, die den Basen -1 bis -540 des Bereiches entspricht, der dem für α -Amylase kodierenden Strukturen vorangeht, oder einem Teil oder der Gesamtheit der DNA-Sequenz, die den Basen -1 bis -320 des Bereichs entspricht, der dem für Glucoamylase kodierenden Strukturen vorangeht.
8. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz einem der Regulons entspricht, die die Expression regulieren von
 - a) ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, wobei das Regulon bevorzugt von Saccharomyces cerevisiae erhalten ist,
 - b) LAC4, wobei das Regulon bevorzugt von Klyveromyces lactis erhalten ist, oder
 - c) korrespondierenden Regulons, bevorzugt von Schwanniomyces erhalten.
9. Verfahren nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß die zweite DNA-Sequenz für die Gesamtheit oder einen Teil mindestens eines der folgenden Peptide kodiert:
 - a) **MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAla;**

b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaAlle
LeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;

5

c) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu Val Ser Ala;

10

d) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu Val Ser Ala Ile Gln Ala Ala Pro Ala;

15

e) Met Ile Phe Leu Lys Leu Ile Lys Ser Ile Val Ile Gly Leu Gly Leu Val Ser Ala Ile Gly Ile Ala Ala Pro Ala Ser Ser Ile Gly Ser Ser Ala Ser Ala

10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß die für das Peptid kodierende DNA-Sequenz einem Teil oder der Gesamtheit einer der folgenden DNA-Sequenzen entspricht:

a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTCAGCAATT
TTAGCATTCTCAAGATTGGTATCTGCT,

3

b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TAGCATTCTCAAGATTGGTATCTGCTAACCGATTATTTTGACATGAGA;

3

c) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTAGTGCT:

4

d) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAATTGGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCC;

e) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCCTTCGATTGGATCTAGTGCTTCA
GCA.

11. Verfahren nach einem der Ansprüche 1 bis 10, dadurch gekennzeichnet, daß die dritte DNA-Sequenz, die für ein Fremdprotein kodiert, eine natürliche oder synthetische DNA-Sequenz ist, die für eine Zellulase, Interleukin, insulinähnlichen Wachstumsfaktor, Interferon, Lymphokin, menschlichen Wachstumsfaktor, Nervenwachstumsfaktor,

Aprotinin, Insulin, Hirudin, Hormone, Blutgerinnungsfaktoren, Hepatitis-B-Oberflächen- oder -Kernantigene, virale oder bakterielle Impfstoffe oder humanen Granulocyten/Makrophagen-Kolonie-stimulierenden Faktor kodiert.

5 12. Verfahren nach einem der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß die vierte DNA-Sequenz von einer Hefe eines Angehörigen der Gattungen Saccharomyces, Pichia, Hansenula, bevorzugt Schwanniomyces, stammt.

10 13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß die vierte DNA-Sequenz einem Teil oder der Gesamtheit des Terminators entspricht, der von den Nukleotiden 1537 bis 1740 des Schwanniomyces- α -Amylasegenes umfaßt ist, oder einem Teil oder der Gesamtheit des Terminators, der von den Nukleotiden 2875 bis 3320 des Schwanniomyces-Glucoamylasegenes umfaßt ist.

15 14. Verfahren nach einem der Ansprüche 1 bis 13, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder dritte und/oder vierte DNA-Sequenz durch Insertion oder Deletion oder Substitution eines oder mehrerer Nukleotide modifiziert worden ist, während die biologische Aktivität der DNA-Sequenz beibehalten oder verbessert wird.

20 15. Verfahren nach einem der Ansprüche 1 bis 14, dadurch gekennzeichnet, daß die Expressionskassette von einem zirkulären oder linearen Vektor getragen wird, der gegebenenfalls ein oder mehrere selektive Markergene umfaßt.

25 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß der Vektor eines der folgenden selektiven Markergene umfaßt:
TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, bevorzugt erhalten von Saccharomyces, Hansenula, Pichia oder Schwanniomyces, AMY1 oder GAM1, bevorzugt erhalten von Schwanniomyces.

30 17. Verfahren nach einem der Ansprüche 15 und 16, dadurch gekennzeichnet, daß der Vektor weiter eine DNA-Sequenz umfaßt, die die autonome Replikation und stabile Beibehaltung des Vektors in dem Wirtsorganismus steuern kann.

35 18. Verfahren nach Anspruch 17 oder 18, dadurch gekennzeichnet, daß die DNA-Sequenz von dem Genom eines Angehörigen der Gattungen Saccharomyces, Pichia, Hansenula oder Kluyveromyces abgeleitet ist, bevorzugt Schwanniomyces.

40 19. Verfahren nach einem der Ansprüche 17 oder 18, dadurch gekennzeichnet, daß die DNA-Sequenz aus den autonom replizierenden Sequenzen ARS, bevorzugt einer ARS, die von dem jeweiligen Wirtsorganismus abgeleitet ist, ausgewählt wird.

45 20. Verfahren nach Anspruch 19, dadurch gekennzeichnet, daß die DNA-Sequenz SwARS1 oder SwARS2 ist, die die autonome Replikation und stabile Beibehaltung des Vektors in Schwanniomyces steuern können, wobei SwARS1 von einem Clal-BamHI-Fragment umfaßt ist, das dem in Figur 2a gezeigten EcoRI-Fragment benachbart ist, und SwARS2 von einem EcoRI-BglII-Fragment umfaßt ist, das in dem in Figur 2a gezeigten EcoRI-Fragment enthalten ist.

50 21. Verfahren nach einem der Ansprüche 17 bis 20, dadurch gekennzeichnet, daß die DNA-Sequenz durch Deletion, Insertion oder Substitution eines oder mehrerer Nukleotide modifiziert worden ist, während ihre biologische Funktion beibehalten oder verbessert wird.

55 22. Verfahren nach einem der Ansprüche 1 bis 21, dadurch gekennzeichnet, daß die Expressionskassette in das Genom integriert ist, gegebenenfalls in mehrfachen Kopien, bevorzugt über homologe Rekombination.

23. Verfahren zum Erzeugen eines Polypeptides, wobei ein Hefewirtsorganismus unter geeigneten Bedingungen kultiviert und das Polypeptid in an sich bekannter Weise gewonnen wird, dadurch gekennzeichnet, daß der Hefewirtsorganismus eine Hefezelle gemäß einem der Ansprüche 1 bis 22 ist.

24. Verfahren nach Anspruch 23, dadurch gekennzeichnet, daß Hefezellen der Art Schwanniomyces occidentalis als ein Wirtsorganismus verwendet werden und bevorzugt auf Stärke, Dextrin, Maltose und/oder Pflanzenbiomasse kultiviert werden.

25. Verfahren nach einem der Ansprüche 23 oder 24, dadurch gekennzeichnet, daß Hefezellen der Gattung Schwanniomyces für die Erzeugung von Einzelzellprotein verwendet werden.

26. Verfahren nach einem der Ansprüche 23 bis 25, dadurch gekennzeichnet, daß das Polypeptid Amylase oder Glycoamylase ist.

27. Verwendung einer Hefezelle gemäß einem der Ansprüche 1 bis 22 zur Transformation und Expression von Fremd-
5 genen.

Revendications

10

Revendications pour les Etats contractants suivants : AT, BE, CH, LI, DE, FR, GB, GR, IT, LU, NL, SE

15

1. Cellule de levure du genre *Schwanniomyces*, **caractérisée en ce que** ladite cellule de levure contient au moins une cassette d'expression comprenant :

20

- a) une première séquence d'ADN servant de régulon,
- b) facultativement, une deuxième séquence d'ADN codant pour un peptide signal
- c) une troisième séquence d'ADN codant pour une protéine étrangère, et
- d) facultativement, une quatrième séquence d'ADN servant de terminateur,

25

les première, deuxième et quatrième séquences d'ADN étant obtenues à partir d'une levure.

30

2. Cellule de levure selon la revendication 1, **caractérisée en ce que** lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'un gène codant pour une enzyme amylolytique.

35

3. Cellule de levure selon la revendication 2, **caractérisée en ce que** ladite enzyme amylolytique est une α -amylase ou la glucoamylase.

40

4. Cellule de levure selon l'une quelconque des revendications 1 à 3, **caractérisée en ce que** lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'une levure d'un élément des genres *Debariomyces*, *Saccharomycopsis*, *Lipomyces*, *Pichia* ou *Saccharomyces*, de préférence *Schwanniomyces*.

45

5. Cellule de levure selon la revendication 4, **caractérisée en ce que** ladite levure est *Schwanniomyces occidentalis*.

50

6. Cellule de levure selon l'une quelconque des revendications 1 à 5, **caractérisée en ce que** ladite première séquence d'ADN est constituée d'un fragment *BglII-Xhol* de 1,8 kb, si elle est obtenue à partir du gène de la α -amylase de *Schwanniomyces occidentalis*, ou d'un fragment *BamHI-PvuII* de 1,3 kb si elle est obtenue à partir du gène de la glucoamylase de *Schwanniomyces occidentalis*.

55

7. Cellule de levure selon la revendication 5, **caractérisée en ce que** ladite première séquence d'ADN est constituée en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -540 de la région précédant le gène structural codant pour la α -amylase, ou en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -320 de la région précédant le gène structural codant pour la glucoamylase.

60

8. Cellule de levure selon l'une quelconque des revendications 1 à 5, **caractérisée en ce que** ladite première séquence d'ADN correspond à l'un des régulons régulant l'expression :

65

- a) de *ADH1*, *ADH2*, *PDC1*, *GAL1/10*, *PGK*, *GAPDH*, le régulon étant de préférence obtenu à partir de *Saccharomyces cerevisiae*,
- b) de *LAC4*, le régulon étant de préférence obtenu à partir de *Kluyveromyces lactis*, ou
- c) de régulons correspondants, obtenus de préférence à partir de *Schwanniomyces*.

70

9. Cellule de levure selon l'une quelconque des revendications 1 à 8, **caractérisée en ce que** ladite deuxième séquence d'ADN code pour tout ou partie d'au moins l'un des peptides suivants :

75

a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaAla
LeuAlaPheSerArgLeuValSerAla;

5

b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaAla
LeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;

10

c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAla;

15

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAlaIleGlnAlaAlaProAla;

20

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAlaIleGlnAlaAlaProAlaSerSerIleGlySerSerAlaSerAla

25

10. Cellule de levure selon la revendication 9, caractérisée en ce que la séquence d'ADN codant pour ledit peptide correspond à une partie ou à la totalité des séquences d'ADN suivantes :

30

a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TTAGCATTCTCAAGATTGGTATCTGCTCAACCGATTATTTTGACATGAGA;

35

b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TAGCATTCTCAAGATTGGTATCTGCTCAACCGATTATTTTGACATGAGA;

40

c) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCT;

45

d) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCC;

50

e) ATGATTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCCCTTCGATTGGATCTAGTGCTTC
GCA.

55

11. Cellule de levure selon l'une quelconque des revendications 1 à 10, caractérisée en ce que ladite troisième séquence d'ADN codant pour une protéine étrangère est une séquence d'ADN naturelle ou synthétique codant

pour une cellulase, une interleukine, un facteur de croissance semblable à l'insuline, un interféron, la lymphokine, un facteur de croissance humain, un facteur de croissance des nerfs, l'aprotinine, l'insuline, l'hirudine, les hormones, les facteurs de coagulation du sang, les antigènes de surface ou de cœur de l'hépatite B, les vaccins viraux ou bactériens, ou le facteur stimulant une colonie de granulocytes/macrophages humain.

5

12. Cellule de levure selon l'une quelconque des revendications 1 à 11, **caractérisée en ce que** ladite quatrième séquence d'ADN provient d'une levure d'un élément des genres *Saccharomyces*, *Pichia*, *Hansenula*, de préférence *Schwanniomyces*.
- 10 13. Cellule de levure selon la revendication 12, **caractérisée en ce que** ladite quatrième séquence d'ADN correspond à une partie ou à la totalité du terminateur constitué des nucléotides 1537 à 1740 du gène de la α -amylase de *Schwanniomyces*, ou à une partie ou à la totalité du terminateur constitué des nucléotides 2875 à 3320 du gène de la glucoamylase de *Schwanniomyces*.
- 15 14. Cellule de levure selon l'une quelconque des revendications 1 à 13, **caractérisée en ce que** lesdites première et/ou deuxième et/ou troisième et/ou quatrième séquences d'ADN ont été modifiées en insérant ou en déletant ou en substituant un ou plusieurs nucléotides tout en retenant ou en améliorant l'activité biologique de la séquence d'ADN.
- 20 15. Cellule de levure selon l'une quelconque des revendications 1 à 14, **caractérisée en ce que** ladite cassette d'expression est portée par un vecteur circulaire ou linéaire, qui comprend facultativement un ou plusieurs gènes de marqueur de sélection.
- 25 16. Cellule de levure selon la revendication 15, **caractérisée en ce que** ledit vecteur comprend l'un quelconque des gènes de marqueur de sélection suivants : TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, obtenus de préférence à partir de *Saccharomyces*, *Hansenula*, *Pichia* ou *Schwanniomyces*, AMY1 ou GAM1, obtenus de préférence à partir de *Schwanniomyces*.
- 30 17. Cellule de levure selon l'une quelconque des revendications 15 à 16, **caractérisée en ce que** ledit vecteur comprend de plus une séquence d'ADN capable de contrôler la réplication autonome et le maintien stable du vecteur dans l'organisme hôte.
- 35 18. Cellule de levure selon la revendication 17, **caractérisée en ce que** ladite séquence d'ADN est obtenue à partir du génome d'un élément des genres *Saccharomyces*, *Pichia*, *Hansenula* ou *Kluyveromyces*, de préférence *Schwanniomyces*.
19. Cellule de levure selon l'une quelconque des revendications 17 et 18, **caractérisée en ce que** ladite séquence d'ADN est choisie parmi les séquences se répliquant de façon autonome ARS, de préférence une ARS obtenue à partir de l'organisme hôte respectif.
- 40 20. Cellule de levure selon la revendication 19, **caractérisée en ce que** ladite séquence d'ADN est SwARS1 ou SwARS2, capable de contrôler la réplication autonome et le maintien stable du vecteur dans *Schwanniomyces*, dans laquelle SwARS1 est constituée d'un fragment *Clal*-*Bam*HI adjacent au fragment *Eco*RI présenté à la figure 2a, et SwARS2 est constituée d'un fragment *Eco*RI-*Bg*II contenu dans le fragment *Eco*RI présenté à la figure 2a.
- 45 21. Cellule de levure selon l'une quelconque des revendications 17 à 20, **caractérisée en ce que** ladite séquence d'ADN a été modifiée en déletant, insérant ou substituant un ou plusieurs nucléotides, tout en retenant ou en améliorant sa fonction biologique.
- 50 22. Cellule de levure selon l'une quelconque des revendications 1 à 21, **caractérisée en ce que** ladite cassette d'expression est intégrée dans le génome, facultativement en plusieurs copies, de préférence par recombinaison homologue.
- 55 23. Procédé pour la préparation d'un polypeptide, dans lequel un organisme hôte de levure est cultivé dans des conditions appropriées et le polypeptide est récupéré d'une manière connue d'elle-même, **caractérisé en ce que** ledit organisme hôte de levure est une cellule de levure selon l'une quelconque des revendications 1 à 22.
24. Procédé selon la revendication 23, **caractérisé en ce que** des cellules de levure de l'espèce *Schwanniomyces*

occidentalis sont utilisées en tant qu'organisme hôte et sont de préférence cultivées sur de l'amidon, de la dextrine, du maltose et/ou de la biomasse végétale.

5 25. Procédé selon l'une quelconque des revendications 23 et 24, **caractérisé en ce que des cellules de levure du genre *Schwanniomyces* sont utilisées pour la production d'une seule protéine de cellule.**

10 26. Procédé selon l'une quelconque des revendications 23 à 25, **caractérisé en ce que l'édit polypeptide est l'amylase ou la glucoamylase.**

15 27. Utilisation d'une cellule de levure selon l'une quelconque des revendications 1 à 22 pour la transformation et l'expression de gènes étrangers.

Revendications pour l'Etat contractant suivant: ES

15 1. Procédé pour préparer une cellule de levure capable d'exprimer une protéine étrangère, **caractérisé en ce qu'au moins une cassette d'expression comprenant :**

20 a) une première séquence d'ADN-servant de régulon,
b) facultativement, une deuxième séquence d'ADN codant pour un peptide signal
c) une troisième séquence d'ADN codant pour une protéine étrangère, et
d) facultativement, une quatrième séquence d'ADN servant de terminateur,
25 est insérée dans une cellule de levure du genre *Schwanniomyces*, les première, deuxième et quatrième séquences d'ADN étant obtenues à partir d'une levure.

30 2. Procédé selon la revendication 1, **caractérisé en ce que lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'un gène codant pour une enzyme amylolytique.**

35 3. Procédé selon la revendication 2, **caractérisé en ce que ladite enzyme amylolytique est une α -amylase ou la glucoamylase.**

40 4. Procédé selon l'une quelconque des revendications 1 à 3, **caractérisé en ce que lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'une levure d'un élément des genres *Debarriomyces*, *Saccharomyces*, *Lipomyces*, *Pichia* ou *Saccharomyces*, de préférence *Schwanniomyces*.**

45 5. Procédé selon la revendication 4, **caractérisé en ce que ladite levure est *Schwanniomyces occidentalis*.**

50 6. Procédé selon l'une quelconque des revendications 1 à 5, **caractérisé en ce que ladite première séquence d'ADN est constituée d'un fragment *Bgl*II-*Xba*II de 1,8 kb, si elle est obtenue à partir du gène de la α -amylase de *Schwanniomyces occidentalis*, ou d'un fragment *Bam*HI-*Pvu*II de 1,3 kb si elle est obtenue à partir du gène de la glucoamylase de *Schwanniomyces occidentalis*.**

55 7. Procédé selon la revendication 5, **caractérisé en ce que ladite première séquence d'ADN est constituée en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -540 de la région précédant le gène structural codant pour la α -amylase, ou en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -320 de la région précédant le gène structural codant pour la glucoamylase.**

60 8. Procédé selon l'une quelconque des revendications 1 à 5, **caractérisé en ce que ladite première séquence d'ADN correspond à l'un des régulons régulant l'expression :**

65 a) de ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, le régulon étant de préférence obtenu à partir de *Saccharomyces cerevisiae*,
b) de LAC4, le régulon étant de préférence obtenu à partir de *Kluyveromyces lactis*, ou
c) de régulons correspondants, obtenus de préférence à partir de *Schwanniomyces*.

70 9. Procédé selon l'une quelconque des revendications 1 à 8, **caractérisé en ce que ladite deuxième séquence d'ADN code pour tout ou partie d'au moins l'un des peptides suivants :**

a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle
LeuAlaPheSerArgLeuValSerAla;

5

b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle
LeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;

10

c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAla;

15

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAlaIleGlnAlaAlaProAla;

20

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu
ValSerAlaIleGlnAlaAlaProAlaSerSerIleGlySerSerAlaSerAla

25

10. Procédé selon la revendication 9, caractérisé en ce que la séquence d'ADN codant pour ledit peptide correspond
à une partie ou à la totalité d'une des séquences d'ADN suivantes :

30

a) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TTAGCATTCTCAAGATTGGTATCTGCT,

35

b) ATGAGATTTCAACTGAAGGATTACAAGTAAAGTTGTTGCAGCAATT
TAGCATTCTCAAGATTGGTATCTGCTCAACCGATTATTTTGACATGAGA;

40

c) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCT;

45

d) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCC;

50

e) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAATTGGTTGGGATTA
GTTAGTGCTATCCAAGCAGCCCCCTGCCCTTCGATTGGATCTAGTGCTTCA
GCA.

11. Procédé selon l'une quelconque des revendications 1 à 10, **caractérisé en ce que** ladite troisième séquence d'ADN codant pour une protéine étrangère est une séquences d'ADN naturelle ou synthétique codant pour une cellulase, une interleukine, un facteur de croissance semblable à l'insuline, un interféron, la lymphokine, un facteur de croissance humain, un facteur de croissance des nerfs, l'aprotinine, l'insuline, l'hirudine, les hormones, les facteurs de coagulation du sang, les antigènes de surface ou de coeur de l'hépatite B, les vaccins viraux ou bactériens, ou le facteur stimulant une colonie de granulocytes/macrophages humain.

5

12. Procédé selon l'une quelconque des revendications 1 à 11, **caractérisé en ce que** ladite quatrième séquence d'ADN provient d'une levure d'un élément des genres *Saccharomyces*, *Pichia*, *Hansenula*, de préférence *Schwan- niomyces*.

10

13. Procédé selon la revendication 12, **caractérisé en ce que** ladite quatrième séquence d'ADN correspond à une partie ou à la totalité du terminateur constitué des nucléotides 1537 à 1740 du gène de la α -amylase de *Schwan- niomyces*, ou à une partie ou à la totalité du terminateur constitué des nucléotides 2875 à 3320 du gène de la glucoamylase de *Schwanniomyces*.

15

14. Procédé selon l'une quelconque des revendications 1 à 13, **caractérisé en ce que** lesdites première et/ou deuxiè- me et/ou troisième et/ou quatrième séquences d'ADN ont été modifiées en insérant ou en déletant ou en substituant un ou plusieurs nucléotides tout en retenant ou en améliorant l'activité biologique de la séquence d'ADN.

20

15. Procédé selon l'une quelconque des revendications 1 à 14, **caractérisé en ce que** ladite cassette d'expression est portée par un vecteur circulaire ou linéaire, qui comprend facultativement un ou plusieurs gènes de marqueur de sélection.

25

16. Procédé selon la revendication 15, **caractérisé en ce que** ledit vecteur comprend l'un quelconque des gènes de marqueur de sélection suivants : TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, obtenus de préférence à partir de *Saccharomyces*, *Han- senula*, *Pichia* ou *Schwanniomyces*, AMY1 ou GAM1, obtenus de préférence à partir de *Schwanniomyces*.

30

17. Procédé selon l'une quelconque des revendications 15 à 16, **caractérisé en ce que** ledit vecteur comprend de plus une séquence d'ADN capable de contrôler la réplication autonome et le maintien stable du vecteur dans l'organisme hôte.

35

18. Procédé selon la revendication 17, **caractérisé en ce que** ladite séquence d'ADN est obtenue à partir du génome d'un élément des genres *Saccharomyces*, *Pichia*, *Hansenula* ou *Kluyveromyces*, de préférence *Schwanniomyces*.

40

19. Procédé selon l'une quelconque des revendications 17 et 18, **caractérisé en ce que** ladite séquence d'ADN est choisie parmi les séquences se répliquant de façon autonome ARS, de préférence une ARS obtenue à partir de l'organisme hôte respectif.

45

20. Procédé selon la revendication 19, **caractérisé en ce que** ladite séquence d'ADN est SwARS1 ou SwARS2, capable de contrôler la réplication autonome et le maintien stable du vecteur dans *Schwanniomyces*, dans laquelle SwARS1 est constituée d'un fragment *Clal-BamH*I adjacent au fragment *EcoRI* présenté à la figure 2a, et SwARS2 est constituée d'un fragment *EcoRI-Bgl*II contenu dans le fragment *EcoRI* présenté à la figure 2a.

21. Procédé selon l'une quelconque des revendications 17 à 20, **caractérisé en ce que** ladite séquence d'ADN a été modifiée en déletant, insérant ou substituant un ou plusieurs nucléotides, tout en retenant ou en améliorant sa fonction biologique.

50

22. Procédé selon l'une quelconque des revendications 1 à 21, **caractérisé en ce que** ladite cassette d'expression est intégrée dans le génome, facultativement en plusieurs copies, de préférence par recombinaison homologue.

23. Procédé pour la préparation d'un polypeptide, dans lequel un organisme hôte de levure est cultivé dans des conditions appropriées et le polypeptide est récupéré d'une manière connue d'elle-même, **caractérisé en ce que** ledit organisme hôte de levure est une cellule de levure selon l'une quelconque des revendications 1 à 22.

55

24. Procédé selon la revendication 23, **caractérisé en ce que** des cellules de levure de l'espèce *Schwanniomyces occidentalis* sont utilisées en tant qu'organisme hôte et sont de préférence cultivées sur de l'amidon, de la dextrine,

du maltose et/ou de la biomasse végétale.

25. Procédé selon l'une quelconque des revendications 23 et 24, **caractérisé en ce que** des cellules de levure du genre *Schwanniomyces* sont utilisées pour la production d'une seule protéine de cellule.

5
26. Procédé selon l'une quelconque des revendications 23 à 25, **caractérisé en ce que** ledit polypeptide est l'amylase ou la glucoamylase.

10
27. Utilisation d'une cellule de levure selon l'une quelconque des revendications 1 à 22 pour la transformation et l'expression de gènes étrangers.

15

20

25

30

35

40

45

50

55

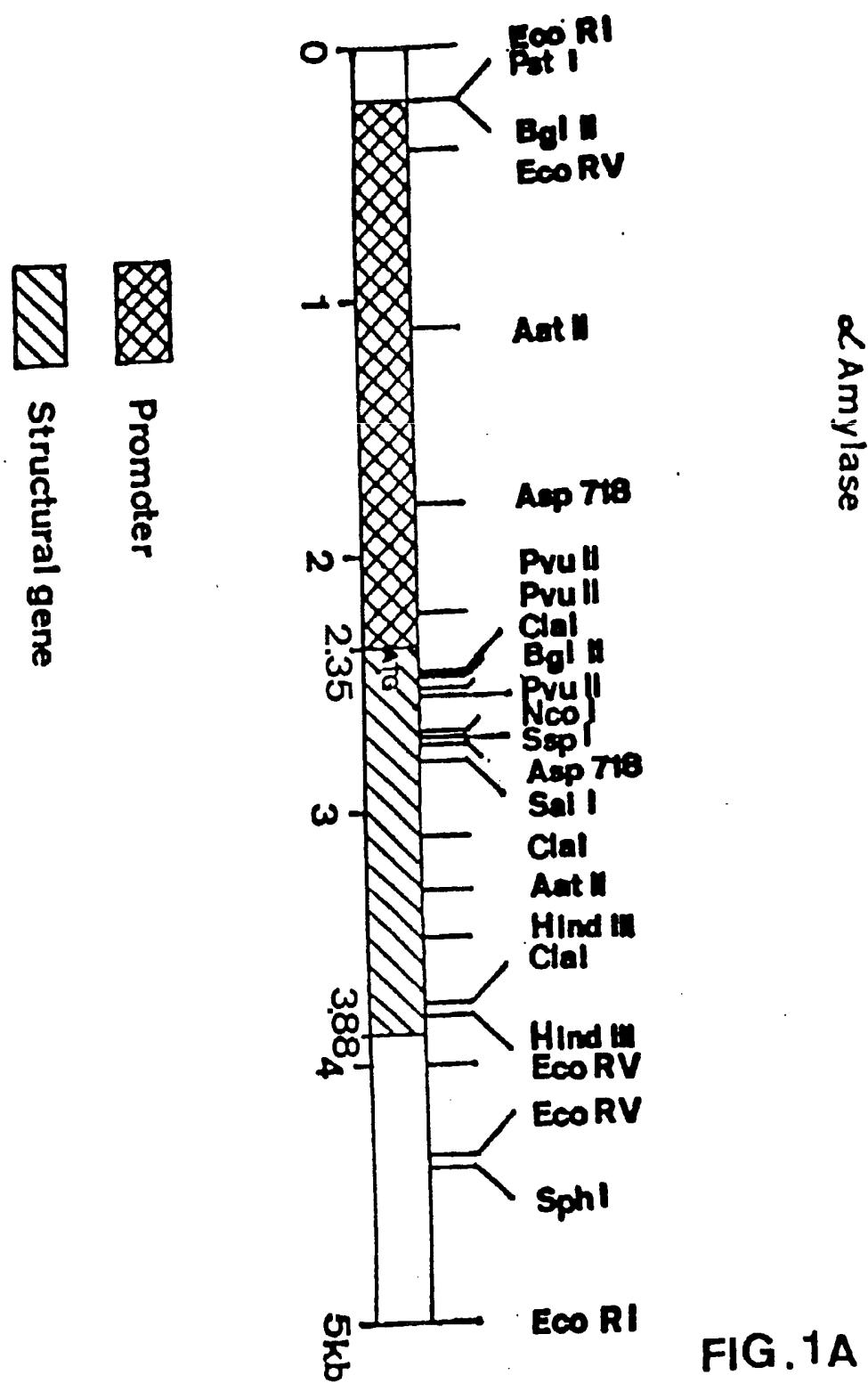


FIG. 1A

FIG. 1B,1

GGTACCTGA

-340 GCTAAATTAA GAACCGCTA TAGATCCGCT TGTCTAAAGA AGAGATAATG AAGAAACAA

-480 TTAACCGAGC ACTCTTATTA AGTTTTTTC TATTTCTTT TGCTCCTACT TCAATAATT

-420 ATCTAAATTG TATTGTGCCT TAGATCAGAA TGTACTGATA ACAGAGAGTA TTATCATACA

-360 CTTGGGATT TCAAAAGGCC GAATCAAAAG CATACTAAT CAAACCCCTTG GTTATTGAT

-300 GCAATTAAAGG TTGTAACTCGT TCTTACCGAT CCATCATTTAT ACCCCACACG GTTTCATGGT

-240 ATGTAGGTGT TTCAATAAGTGT AGTACAATG AATGTTTTGG TAATGCTGTA TGTGGATCAG

-180 TAATTATGTT AAACAATTAA GTCTGAAAT TTATTAAT TTACCTACA ATTAAGCCG

-120 AAATCCAATC GAAGGTGCCG CCCAGCTGGT GTATAAATTA CTCTTGAAT TCAAGTTGAA

-60 CGTTGATCTC TCTAAAGCAA AGCTGTATT CTACATTAATC AAATAAAATA AAAGCAAGAC

10 20 30 40 50 60
 ATGAGATTTCACACTGAAAGGATTACAAGTAAAGTTGTTGCAGGCAATTTCAGCATTCTCA
 Met Arg Phe Ser Thr Glu Glu Ile Phe Thr Ser Lys Val Val Ala Ile Leu Ala Phe Ser

 70 80 90 100 110 120
 AGATTGGTATCTGCTCAACCGATTATTGACATGAGAGATGTTAGCTCGTCAGCTGAT
 Arg Leu Val Ser Ala Gln Pro Ile Ile Phe Asp Phe Asp Met Arg Asp Val Ser Ser Ala Asp
 ★

 130 140 150 160 170 180
 AAATGGAAAGACCAATCGATTATCAAATCGTTACTGATAAGGTTGCCAGATCTGATGGC
 Lys Trp Lys Asp Gln Ser Ile Tyr Gln Ile Val Thr Asp Arg Phe Ala Arg Ser Asp Gly
 ★

 190 200 210 220 230 240
 TCGACCACAGCTGACTGTTAGTGAGTGATCGCAAGTACTGTGGTGGATCTTATAAAGGG
 Ser Thr Thr Ala Asp Cys Leu Val Ser Asp Arg Gly Lys Tyr Cys Gln Ile Val Ser Tyr Lys Gly
 ★

 250 260 270 280 290 300
 ATTATCGACAAGTTGGATTATTCAAGGTATGGGTACTGCGATCTGGATCTCCCCA
 Ile Ile Asp Lys Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala Ile Trp Ile Ser Pro

1b,2

310 320 330 340 350 360
 GTTGTGAGCAAATTCCCTGACAATACTGCCTTATGGTTATGCCTTACCATGGTTATTGGATG
 ValValGluGinIleProAspProAsp
 370 380 390 400 410 420
 AAAATATTGATGAAATTGAAACACTAAATTGGTACCGGTGATGAATTGAAACAAATTAGCT
 LysAsnIleAspGluLeuAsnThrAsnPheGlyThrAlaAspGluLeuLysGinLeuAla
 430 440 450 460 470 480
 AGCGAATTGCATTCCAGAAGCGATGTTATTGATGGTCGACGTTACACCAATTATGCT
 SerGluLeuHisSerArgSerMetLeuMetValAspValValTyrAsnHistYrAla
 490 500 510 520 530 540
 TGGAACGGAGATGGTTCAAGCGTAGATTCTAGTTTCAACTCCATTCAACAAATCT
 TrPAsnGlyAspGlySerSerValAspTyrSerSerPheThrProPheAsnGlnSer
 550 560 570 580 590 600
 TACTTCCACGATTATTGTTGATTACAAATTATAATGATCAAAACCAATTGTTGAAGATTGT
 TyrPheHisAspTyrCysLeuIleThrAsnTyrAsnAspGlnThrAsnValGluAspCys
 610 620 630 640 650 660
 TGGGAAGGTGATACTGAAGTACCTCCCTCAGATTAAAGTACCGAGGATAATGAAGTTATA
 TrpGluGlyAspThrGluValSerLeuProAspLeuSerThrGluAspAsnGluValIle

1b,3

670 680 690 700 710 720
 GGAGTATTCAAACCTGGGTTCAGATTTCAAACTATTCAATCGATGGTTTAAGA
 GlyValPheGlnThrTrpValSerAspPheAlaGlnAsnTyrSerIleAspGlyIleArg
 730 740 750 760 770 780
 ATTGATAGTGCAAAAGCACGCTAGATAACCGCTTCATTAAACGAAGTTGAGGACGCTCTGGT
 IleAspSerAlaLysHisValAspThrAlaSerLeuThrLysPheGluAspAlaSerGly
 790 800 810 820 830 840
 GTTTAACTTAGGTGAAGTTTATCAAAGGAGATCCAACTTATACTTGTCATATCAGAAT
 ValTyrAsnLeuGlyGluValTyrGlnGlyAspProThrTyrThrCysProThrGlnAsn
 850 860 870 880 890 900
 TATATGAAAGGAGTTACCAACTTACATTACGTTATAGATTCTTCAGTGTAT
 TyrMetLysGlyValThrAsnTyrProLeuThrTyrProValTyrArgPheAspSerAsp
 910 920 930 940 950 960
 ACTTCGGCGACTTCCAGTTAACATGATCTCAATGAGTCCACGTACAGTCATCTGTTCG
 ThrSerAlaThrSerSerGluLeuThrSerMetIleSerThrLeuGlnSerSerCysSer
 970 980 990 1000 1010 1020
 GACGTCTTGTGGAAACTTTATTGAAACCATGACCAAGTTAGATTCCATCAGTT
 SerLeuGlyAsnPheIleGluAsnHisAspGlnValArgPheProSerVal
 1030 1040 1050 1060 1070 1080
 ACCTCAGACACATCCTGATTAAAGAATGACATGGCTTTATAATTGGGTGATGGTATC
 ThrSerAspThrSerSerLeuIleLysAsnAspMetAlaPheIleIleLeuGlyAspGlyIle

1b, 4

1090 1100 1110 1120 1130 1140
 CCAATTATTATGGCCAAAGAACAAAGGTCTCAATGGTGGTTCCGATCCTGCCAATAGA
 ProIleIleTyrTyrGlyGlnGluGlyLeuAsnGlySerAsnProAlaAsnArg
 1150 1160 1170 1180 1190 1200
 GAAGCTTTATGGTTAAGTGGATATAACCGATTAGAAATACTACCGAGCTAATCAGTAAAG
 GluAlaLeuIleSerGlyTyrAsnThrAspSerGluTyrTyrGluLeuIleSerLys
 1210 1220 1230 1240 1250 1260
 CTAATCAAATAAGAAATCAAGCTATTAGAAAGGATTCTGCCTTCAACTTACAAATCC
 LeuAsnGlnIleArgAsnGlnAlaIleLysAspSerAlaTyrSerThrTyrLysSer
 1270 1280 1290 1300 1310 1320
 TCAGTTGGTTCTTCAGACCACATAAGCCACTAGGAAAGGGTAGGGATGCTAAATCAA
 SerValAlSerSerAspHistYrIleAlaThrArgLysGlySerAspAlaAsnGln
 1330 1340 1350 1360 1370 1380
 TGATTTCCATTAAATTAGGGTTCAAACGGGTCAACAGGATATTACTGTCAGCAAC
 LeuIleSerIlePheAsnAsnLeuGlySerAsnGlySerGlnAspIleThrValSerAsn
 1390 1400 1410 1420 1430 1440
 ACCGGCTATTCTAGTGGTGTATAAGTTATCGATATTCTTGCAATTCCGGTTAGCT
 ThrGlyTyrSerSerGlyAspLysValIleAspIleSerCysAsnSerValLeuAla
 1450 1460 1470 1480 1490 1500
 GGTGACTCCGGAAGGCTTATCTGTATCAATTTCGGTGGAAATGCCACAAAGTTACGCTCCG
 GlyAspSerGlySerValSerValSerLeuSerValSerGlyGlyMetProGlnValTyrAlaPro

1b,5

1b,6

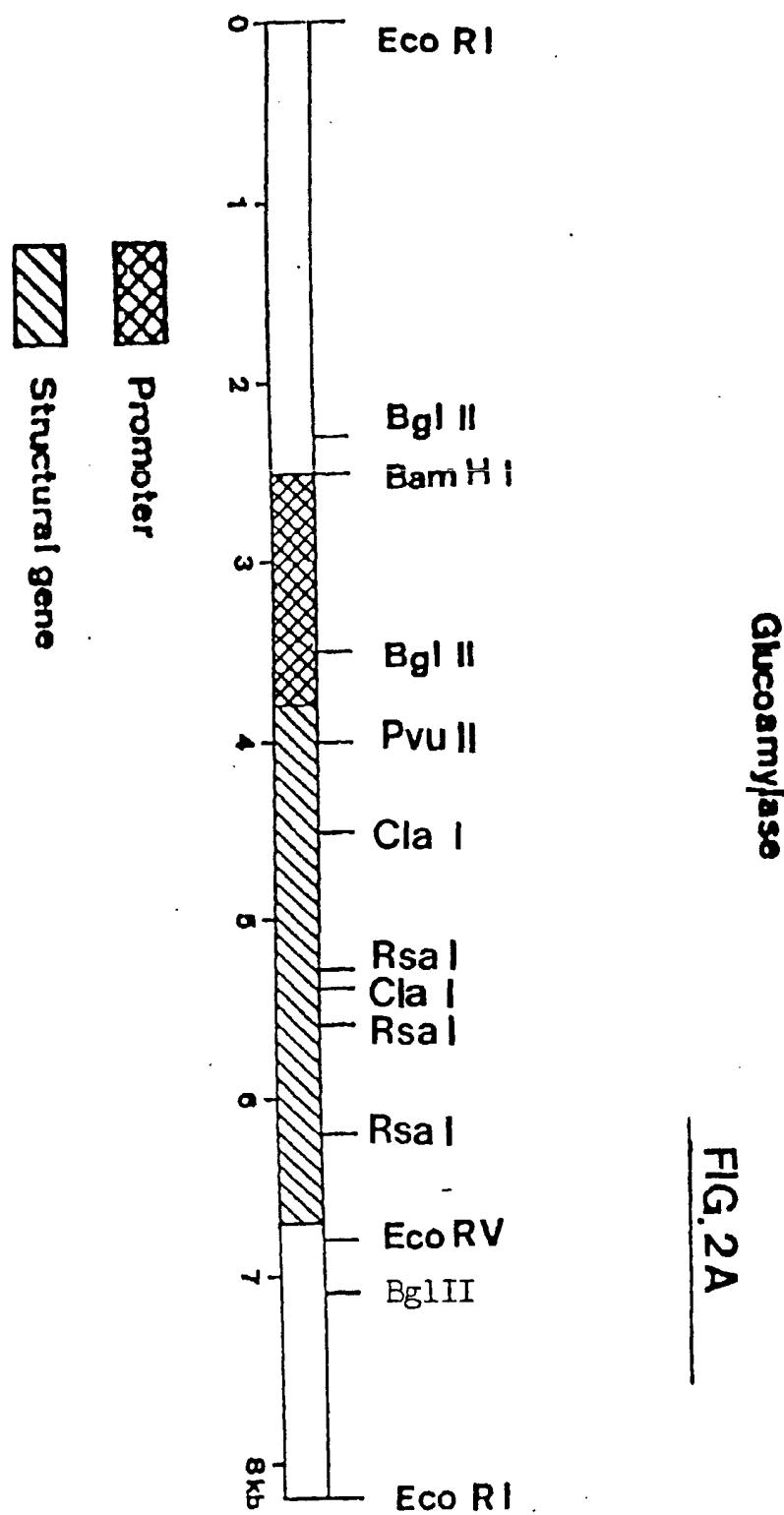


FIG. 2B₁

EP 0 394 538 B1

-310
AGATCTACATTAAACCCCAAG

-250
TCTACTCCAGATATTGGAGTATTAACCCCATTTACCGTTATATCCATGACCCGCGATCGA

-200
AATTTCAAAAGGATTTCAGGAAATTCTTCTAAATACGAAAGTGTATTGGTGTCA

-150
ATTACTACGGAAACTCATATGGTAGAGTTGGTGAATGAAATTAGCGCAATTGGTAAATT

-100
GCGAAGTTATAAGTTGGCAAACCTGGAGAATTTCATTATTGGAAATAATAAA

-50
TAAGGGCAAGTATCCATTGAAATTTAAATGAACTCATGACTGTATATAACAAGCAG

-1

10 20 30 40 50 60
 ATGATTCTGAAAGCTGATTAAAAGTATAAGTAAATTGGTTGGGATTAGTTAGTTAGCTATC
 Met Ile Phe Leu Lys Ser Ile Val Ile Gly Leu Gly Leu Val Ser Ala Ile
 ★
 70 80 90 100 110 120
 CAAGCAGCCCCCTGCCTCTCGATTGGATCTAGTGCCTCAGGCATCTAGTTCAAGTGAAGT
 Gln Ile Ala Pro Ala Ser Ser Ile Gly Ser Ser Ile Ser Ser Gln Ser
 ★
 130 140 150 160 170 180
 TCTCAGGGCTACAAATTCCCAATGATGATACATTAGGTGTTAACAAATTCCCTAAATATCTT
 Ser Gln Ala Thr Ile Pro Asn Asp Val Ile Thr Leu Gly Val Ile Pro Asn Ile Phe
 ★
 190 200 210 220 230 240
 AATGACTCTGCTGCGATGCCTAAATGCAGCTGCTAAAGGGTATGACTGGTAAATGGTTACT
 Asn Asp Ser Ala Val Ile Asp Ala Asn Ala Asn Ala Asn Ala Asn Val Thr
 ★
 250 260 270 280 290 300
 AATACTCCAAAGAGGATTAAACGGTATCTAAATTAAGAAAGCTACCAATATTATGGT
 Asn Thr Pro Arg Gly Ile Leu Thr Gly Ile Leu Ile Leu Ile Val Gly
 ★
 310 320 330 340 350 360
 TATGATTCTGATTAACTTAACCTGAAATACCAAGCTGATAACCAAGATTAAACGTT
 Tyr Asp Phe Asp Tyr Leu Asn Leu Thr Val Gly Ile Tyr Gln Ile Asp Thr Arg Leu Asn Val
 ★
 370 380 390 400 410 420
 CATATTGAACCAACTGATTCTGATGTATTGGTTTACCAAGGGCATTAGTTAGTTGTTAA
 His Ile Gln Pro Thr Asp Val Phe Val Leu Pro Gln Ile Val Val Ile Val Ile

2b, 2

430 440 450 460 470 480
 CCACTGGTGAAGGTGATGCCAACATTCTTAACTTCGACAAATTCCGATTGGTTTCGAA
 ProLeuValGluGlyAspAlaGinSerTyrAsnProAspAsnSerAspLeuValIPheGlu
 490 500 510 520 530 540
 TACTCTAATACTGACTTCTCCCTTTGAAGTTATTAGATCATCTACTAAAGAAGTTTATT
 TyrSerAsnThrAspPheSerPheSerSerPheGluValIleSerSerThrLysGluValLeuPhe
 550 560 570 580 590 600
 TCTACTAAAGGTAATCCATTGGTTTCAAAATTCATTCAATTCAATTTCGTCATG
 SerThrLysGlyAsnProLeuValPheSerAsnGlnPheSerSerSerSerLeu
 610 620 630 640 650 660
 CCAAAAGAACCATGGTATTACTGGTCTTGGTGAATCTTACGGTTAGTTAACCGAACCA
 ProLysAsnHisValIleThrGlyGluSerIleHisGlyLeuValAsnGluPro
 670 680 690 700 710 720
 GGTAGCGGTAAACATTATTGCTTAATGATGTTGGTGAATCGATTGGTAATATTAT
 GlySerValIlyThrLeuPheAlaAsnAspValGlyAspProIleAspGlyAsnIleTyr
 730 740 750 760 770 780
 GGTGTCCCATCCAGTTTATGATCAAAGATATGACACTGAAACTACCCATGCTGTTAT
 GlyValHisProValTyrLeuAspGlnArgTyrAspThrGluThrThrHisAlaValTyr
 790 800 810 820 830 840
 TGCGAGAACCTCTGCTTCAAGAAGTATTCAAGGAAATCTTACTTGGAGAGCT
 TrpArgThrSerAlaIleGlnGluValLeuGlyGluSerIleThrTrpArgAla

2b,3

850 860 870 880 890 900
 CTTTCAGGTATTGATTATACTTCTTAGTGGTCCCTACACCAAAAGATGCCATTCAA
 LeuSerGlyValIleAspLeuTyrPheSerGlyProThrProLysAspAlaIleGln
 910 920 930 940 950 960
 CAGTATGTCAAAGAGATTGGTTACCAACCATACTGGTCTGGTTAGGTTACCAT
 GlnTyrValIleGlyIleGlyLeuProAlaPheGlnProTyrTrpSerLeuGlyTyrHis
 970 980 990 1000 1010 1020
 CAATGAGATGGGGTTACGATACTATCGAAAATTATCTGAAGTTGGAAAACCTTCAG
 GlnCysArgTrpGlyTyrAspThrIleGlyLeuSerGluValIleGlyAsnPhsLys
 1030 1040 1050 1060 1070 1080
 AAATTAAATTCCATTAGAAACTATCTGGTCAGACATTTGATTACATGGACTCTTATAAA
 LysPheAsnIleProLeuGluThrIleTrpSerAspIleAspTyrMetAspSerTyrLys
 1090 1100 1110 1120 1130 1140
 GATTTCACTTATGATCCACAGATTCCCACTAGATGAATACTGTAATTTCCTTGATGAG
 AspPheThrTyrAspProHisArgPheProLysAspGluIleTyrValGlyAspGlu
 1150 1160 1170 1180 1190 1200
 TTGCACAAATAATCAACACTATGTTCTTGGATGCTGCTATTACGTTCCAAAC
 LeuHisLysAsnAsnGlnHistYrValIleProLysLeuAspAlaIleTyrValProAsn
 1210 1220 1230 1240 1250 1260
 CCAACAAATGCTACGGATAACGAAATACCAACCTTTCCACTATGGTAAAGAACCGATGTC
 ProAsnAsnAlaThrAspAsnGluIleTyrGlyAsnProHisTyrGlyAsnGluThrAspVal

2b,4

1270 1280 1290 1300 1310 1320
 TTCTTAAAGAAATCCAGATGGTTCAATTATTTGGCTGGGTTGGGTTACACTGTT
 PheLeuLysAsnProAspGlySerLeuTyrIleGlyAlaValTrpGlnValThrLeuPhe
 1330 1340 1350 1360 1370 1380
 TCCAGATTCTTAGGAGAAACATTCAAGATATGGATAAGTCATTAAAGATTGGTATGAA
 SerArgPheLeuSerArgLysHisSerAspMetAspLysValIleAspTrpTyrGlu
 1390 1400 1410 1420 1430 1440
 TTAACCTCTTGTGGTATTGGCTGATAATGAATGAABTCATCATTCATGTTGGT
 LeuThrProPheAspGlyIleTrpAlaAspMetAsnGluValSerPheCysValGly
 1450 1460 1470 1480 1490 1500
 TCTTGTTGTTACTGGTAATAACTTCGAAACCCAGCATACTCCCATTTACTGTTGGAAAGT
 SerCysGlyThrGlyLysTyrPheGluAsnProAlaTyrProProPheThrValGlySer
 1510 1520 1530 1540 1550 1560
 AAAGCTACCTCTTATCCAGTTGGTTTCGATGTTCTAACGGCATCTGAATGGAAATCTATT
 LysAlaThrSerTyrProValGlyPheAspValSerAsnAlaSerGluTrpLysSerIle
 1570 1580 1590 1600 1610 1620
 CAAAGCTCAATTCTGCTACTGCTAAGACTTCAACTTCTCAACTTCTCCGTTATCCTGCTCTCA
 GlnSerSerIleSerAlaThrAlaLysThrSerSerSerSerSerSerSerSerSer
 1630 1640 1650 1660 1670 1680
 TCCACAAATCGATTATGAAACACTTCTGCTCCAGGTAAAGGTAATTAATTATCCACCA
 SerThrIleAspTyrMetAsnThrLeuAlaProGlyLysGlyAsnIleAsnTyrProPro

2b,5

2b,6

2110 2120 2130 2140 2150 2160
 GCT ATT GAT CAG GAA CCT ATG TGG AAT CAG TGT GCT GAA GCT TAC TGA AACT GAT
 Ala Ile Asp Gln Gln Pro Tyr Val Trp Glu Ser Val Ala Glu Val Thr Arg Thr Ser Met

2170 2180 2190 2200 2210 2220
 GCC ATT AGA TCT ATT TAC CTT TAC ATT TAC ATG AA TCT CTA CTA CTA CTA CTC
 Ala Ile Arg Tyr Lys Leu Leu Pro Tyr Tyr Tyr Thr Leu Lys Glu Ser His Thr Thr

2230 2240 2250 2260 2270 2280
 GGT TAC CAA ATC TTA AGA GCT TCT CGT GGC AA ATC CGT CTC GTC CTC CTC TAA GTC
 Gly Leu Pro Ile Leu Lys Arg Ala Phe Ser Trp Glu Pro Asn Asp Arg Ser Leu Ser Gly

2290 2300 2310 2320 2330 2340
 GTC GATA ACC AA TTT TGT GTC GGG TTA GAT GGT TTA GGT TAA TCT CTC GTC TCT TAA GAC
 Val Asp Asn Gln Phe Phe Val Phe Val Val Thr Pro Val Leu Glu Pro Gly Val

2350 2360 2370 2380 2390 2400
 GTG GATA AGG GT TAA AGG GT TAA AGG GAG CTG GT TAA AGG GAA AGT TAA TCT AC GACT GG
 Val Asp Lys Gln Val Phe Pro Gln Val Gln Val Val Val Thr Pro Val Leu Glu Pro Gly Val

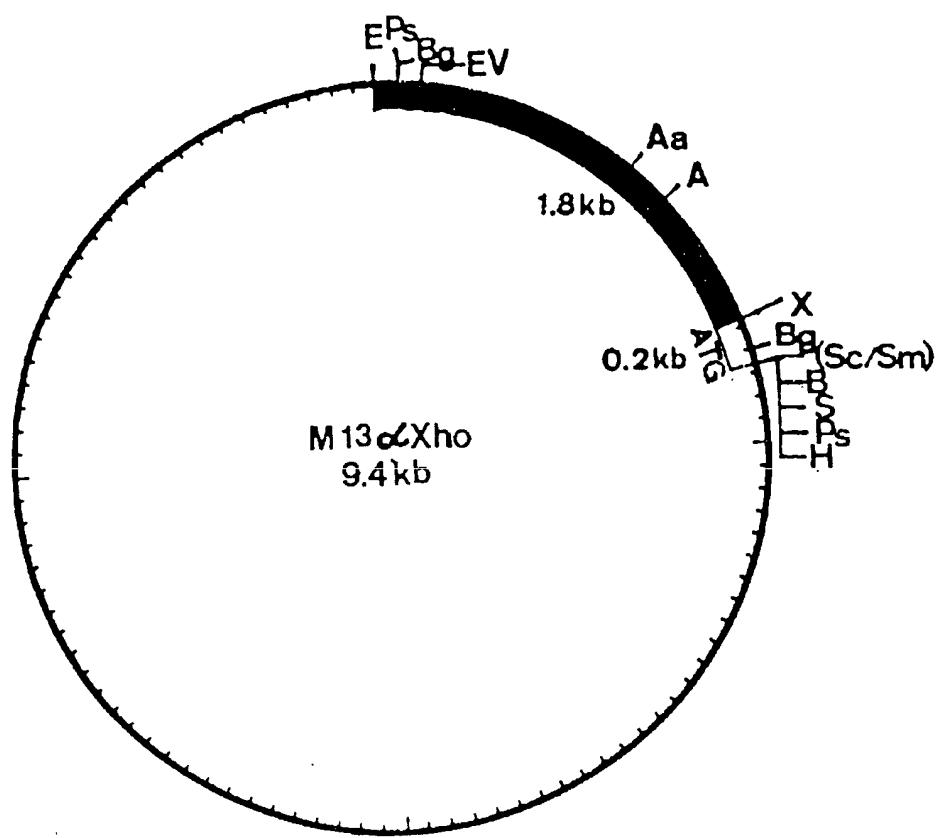
2410 2420 2430 2440 2450 2460
 TAC ACC CA AAG GAA GGT TCA CTT TAA AGA GCT GGT TAA GAA ATG AA CTT TAA GAT GCA CCA CTT A
 Tyr Thr Glu Arg Glu Val His Phe Lys Asp Gln Lys Asn Glu Thr Leu Asp Ala Pro Leu

2470 2480 2490 2500 2510 2520
 GGT CATA TCC ATT TAC A CTT A GGT GGT AAC GT TGG CAA ACT CAA GAG CC AAC GCT TAA
 Gly His Ile Pro Leu His Ile Arg Gln Val Val Pro Thr Glu Glu Pro Gly Tyr

2b,8

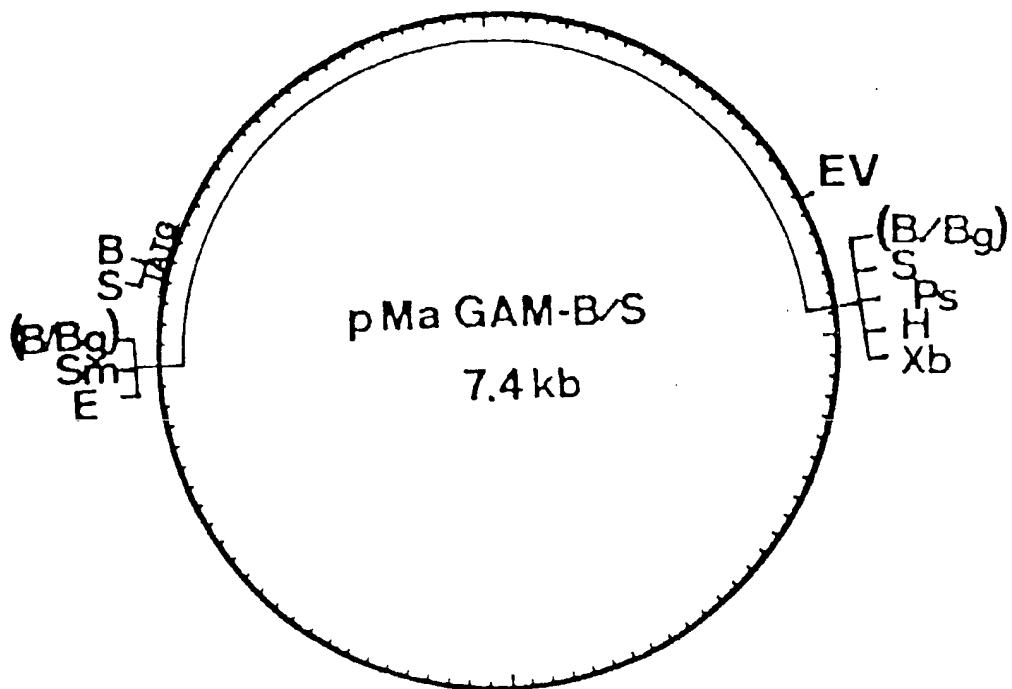
2950 2960 2970 2980 2990 3000
TTATTTGAATGATATCATTTAGTTTCTTCAGCTTATGCTATGTTATTGATTAGTAAATAGTTAA
ATGTTAAAAGTTTTTATGTTATGTTATTGATTAGTTGATGTTGATTAAGGAA
3010 3020 3030 3040 3050 3060
ATGTTAAAAGTTTTTATGTTATGTTATTGATTAGTTGATGTTGATTAAGGAA
3070 3080 3090 3100 3110 3120
TACATCGAAAATTGGATGAAATTGGATGAAATTGGATGAAATTGGATGAAATGTTAGTCATTAGT
3130 3140 3150 3160 3170 3180
CACATTGGACCACTATAACAAACAAACTATTCCAAAGAAAAATATATGTAAGGATAC
3190 3200 3210 3220 3230 3240
TAGATCATAAATTCTTATTGACTTGTGTTTAAACAATAGTTACATAAGGAATATTTCG
3250 3260 3270 3280 3290 3300
TTTACTACAAACCATGGTCTGTAAAGAAGGCAGACGAGGGCTATGTTGGTTGGGG
3310 3320
CGCGCAATACTAGTTACAAAG

2b,9



—	M13	Aa	Aat II
□	α Amylase 0.2kb of the structural gene	A	Asp 718
■	α Amylase-Promoter	B	Bam H I
E	Eco R I	Sm	Sma I
Ps	Pst I	S	Sal I
B _g	Bgl II	H	Hind III
EV	Eco RV	Sc	Scal I
		X	Xho I

FIG. 3



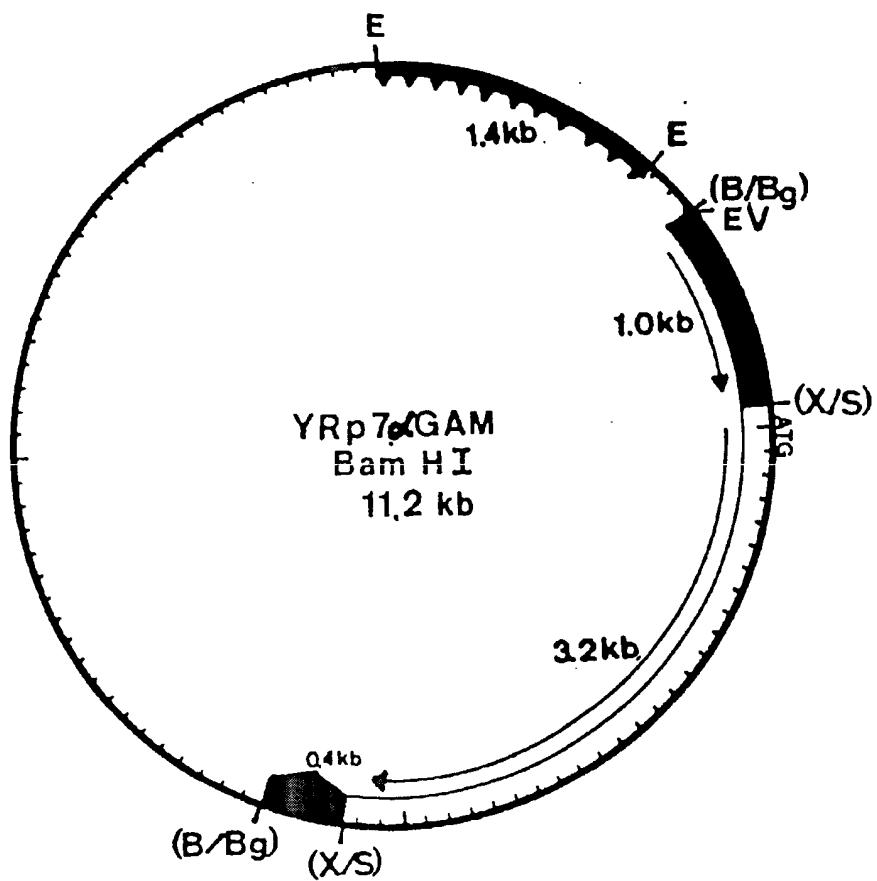
□ Glucoamylase

— pMac5-8
Kramer et al. 1984

EV Eco RV
B Bam HI
Bg Bgl II
S Sal I
Ps Pst I

H Hind III
Xb Xba I
E Eco RI
Sm Sma I

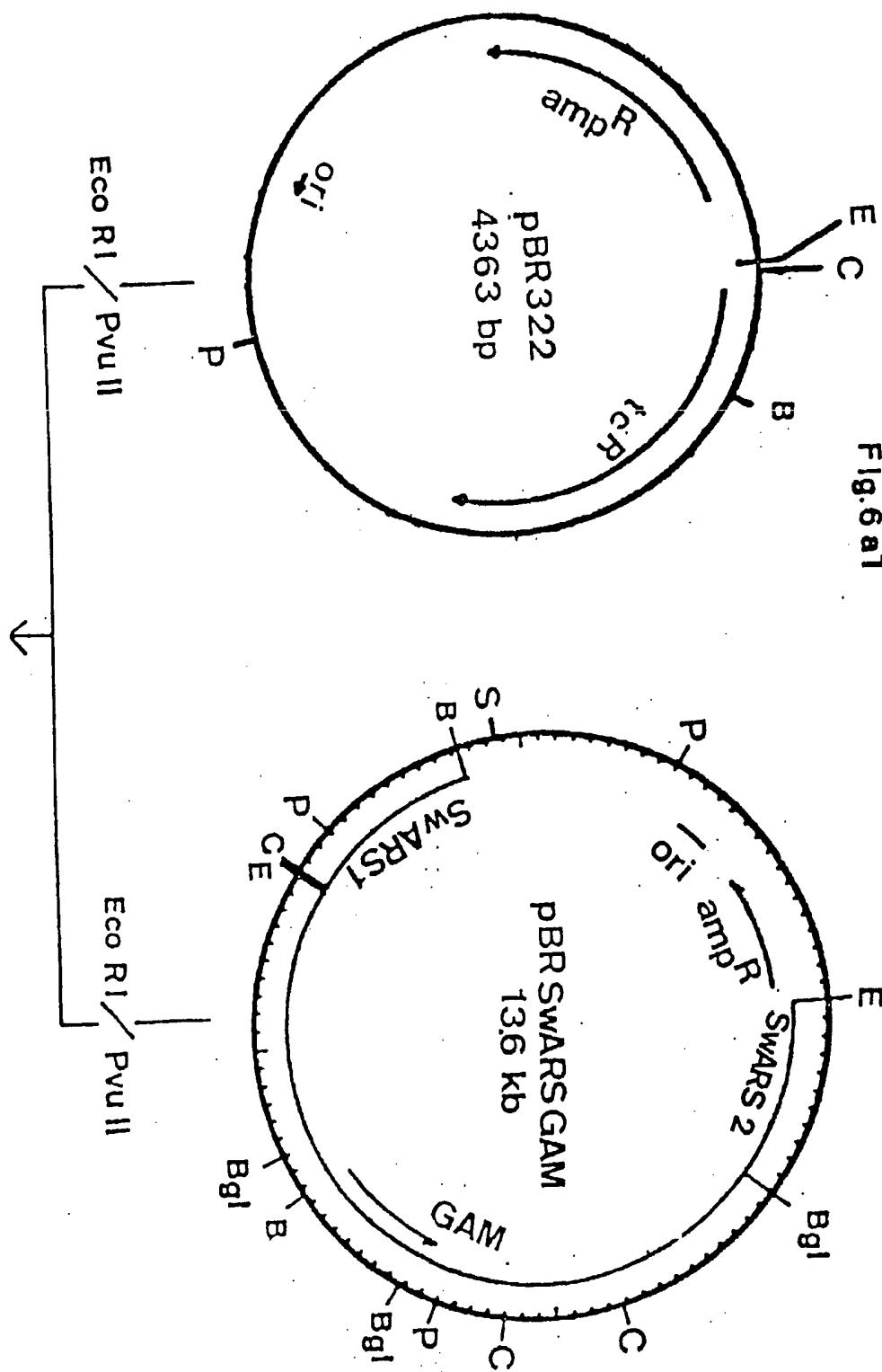
FIG.4



■■■	TRP/ ARS	E	Eco R I
■■■	α Amylase-Promoter	B	Bam H I
□□□	Glucoamylase structural gene	Bg	Bgl II
—	YRp7 Struhl et al. 1979	EV	Eco R V
■■■	α Amylase 0.4 kb of the structural gene	X	Xho I
		S	Sal I

FIG.5

Fig. 6 a1



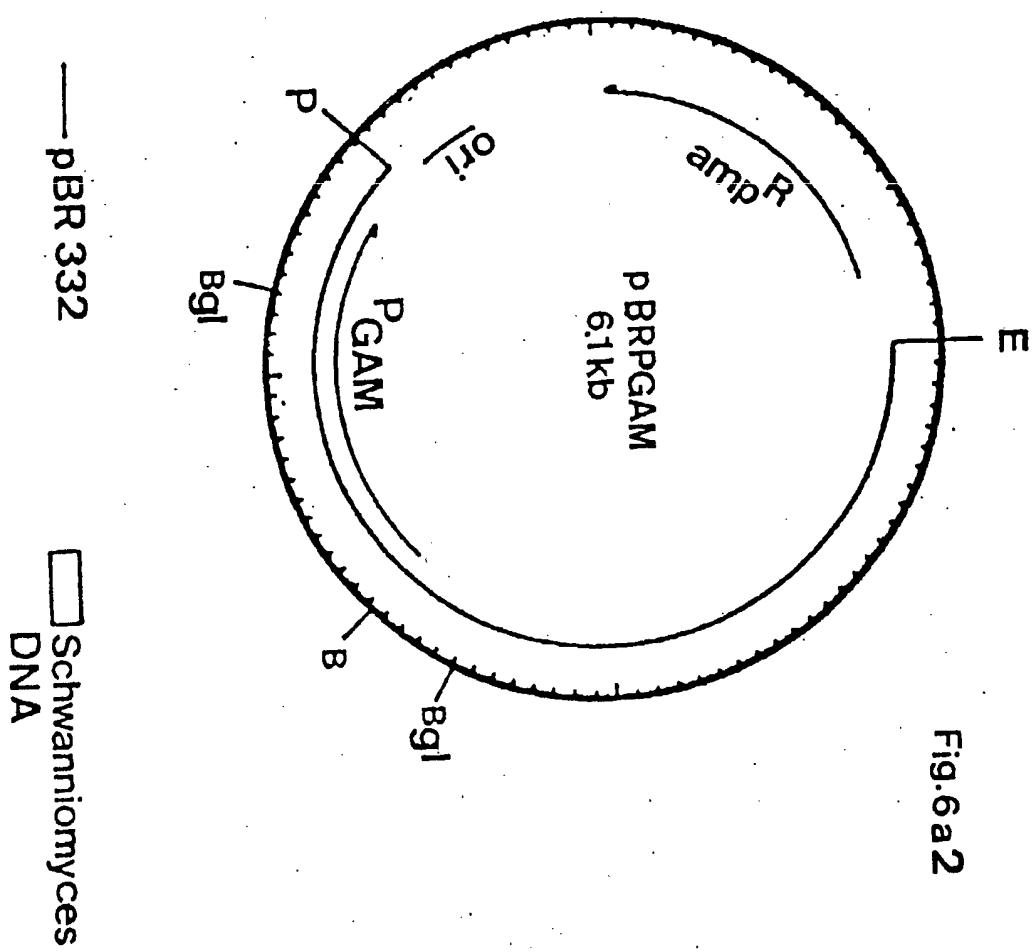


Fig.6 a2

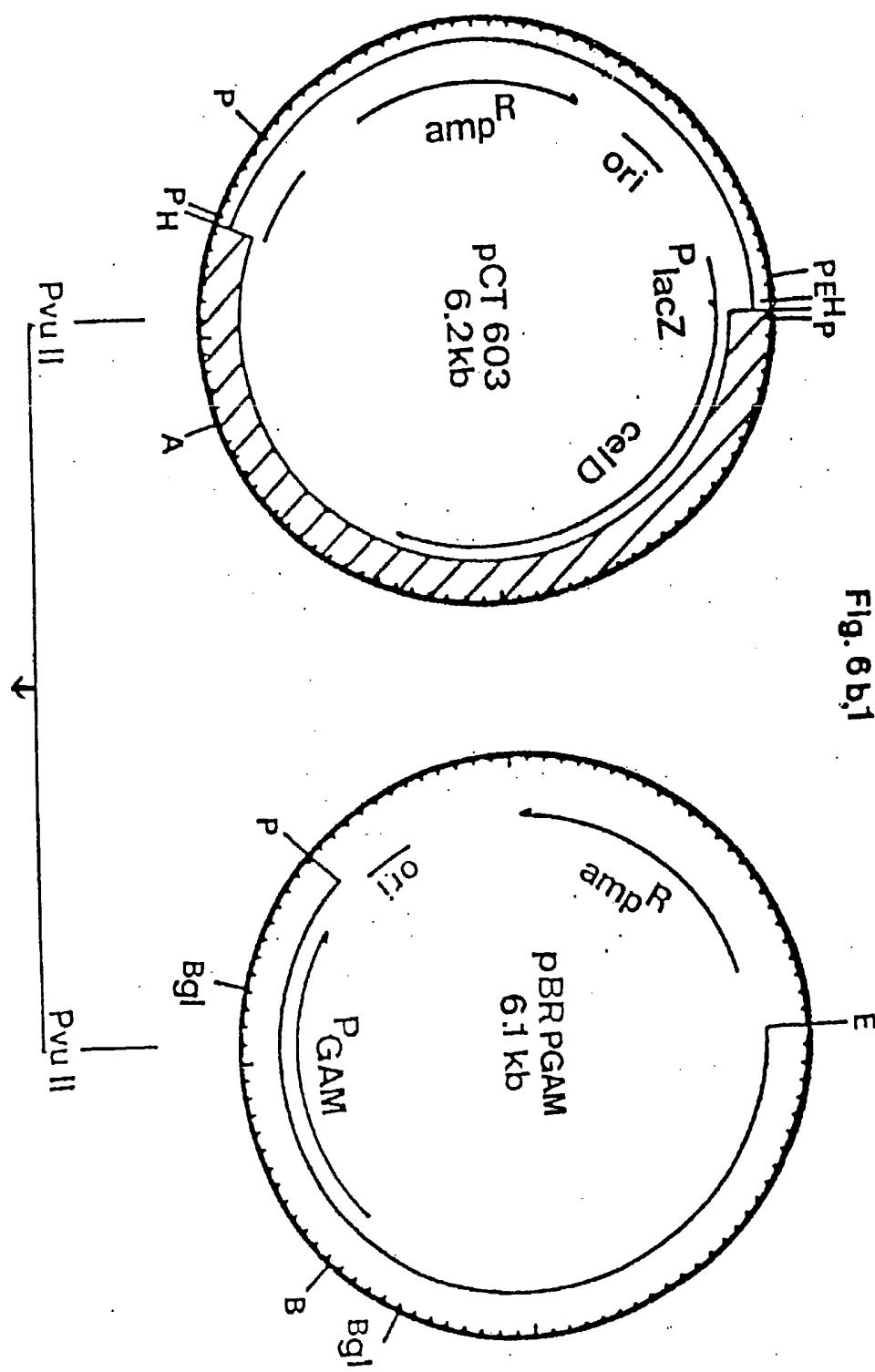
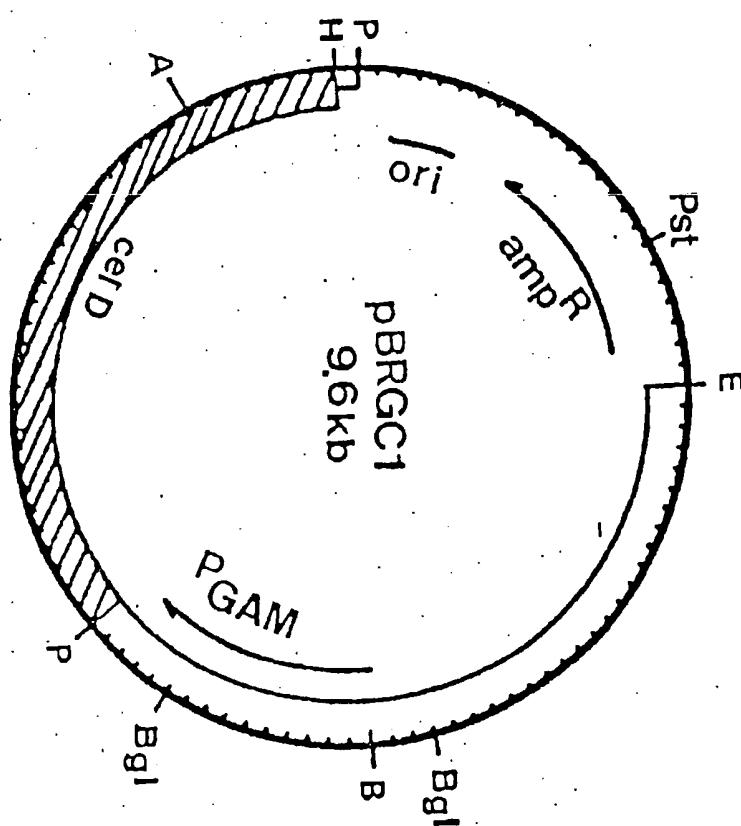


Fig. 6 b,1

Fig.6b,2



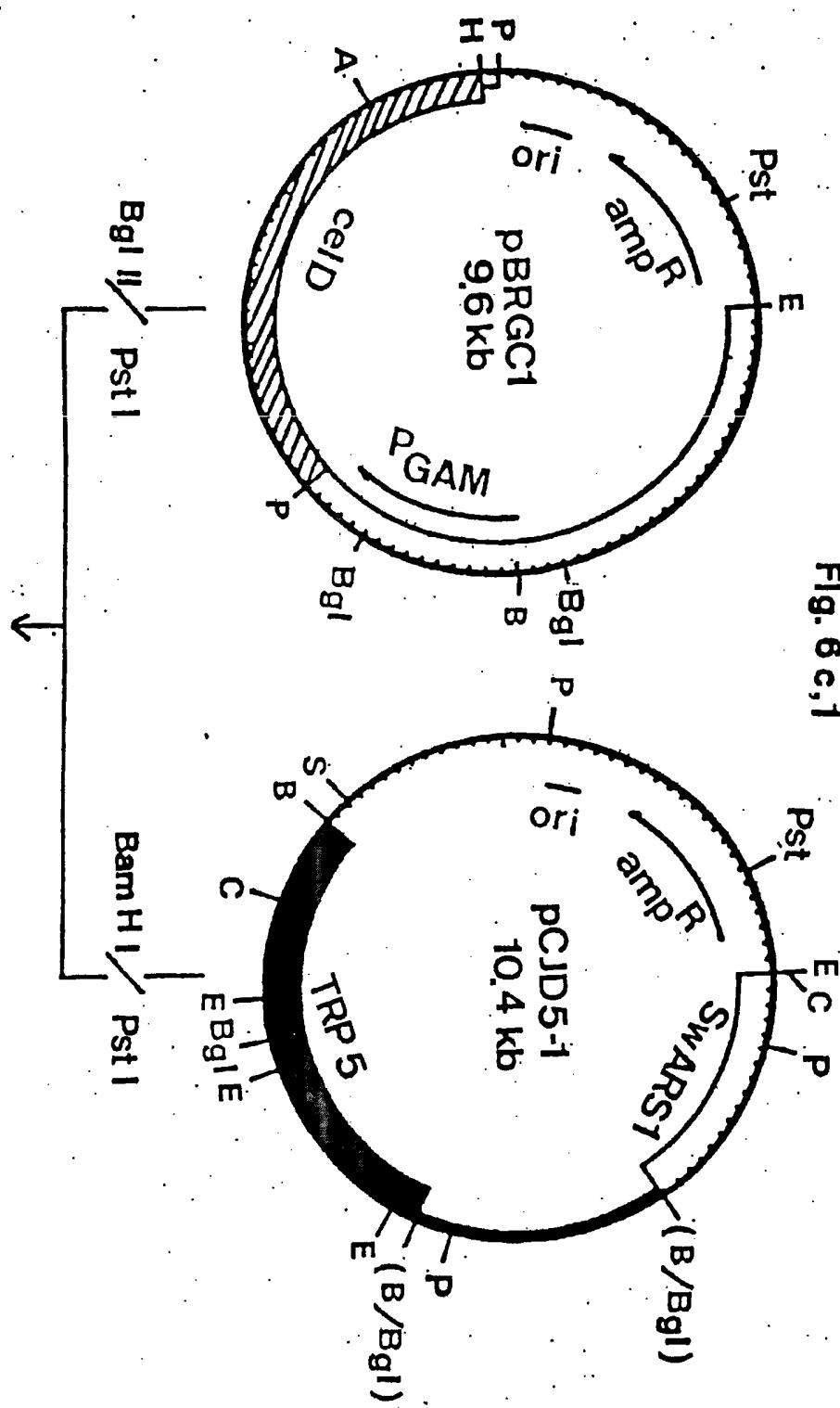
— pBR 332

≡ pUC 8

▨ C.thermocellum DNA

◻ Schwanniomyces DNA

Fig. 6 c,1



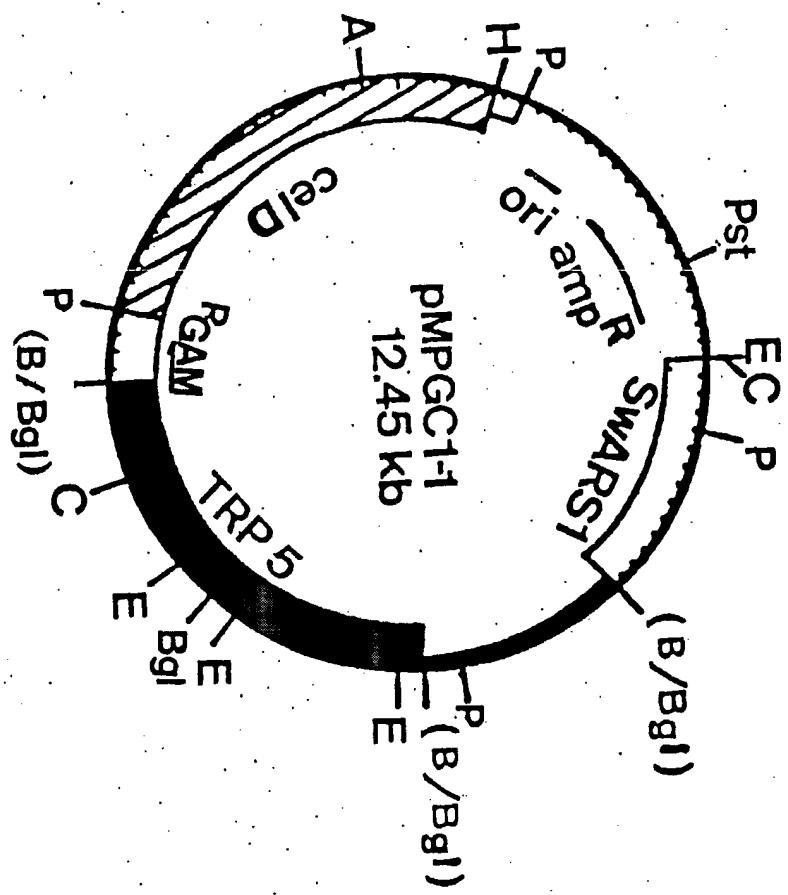
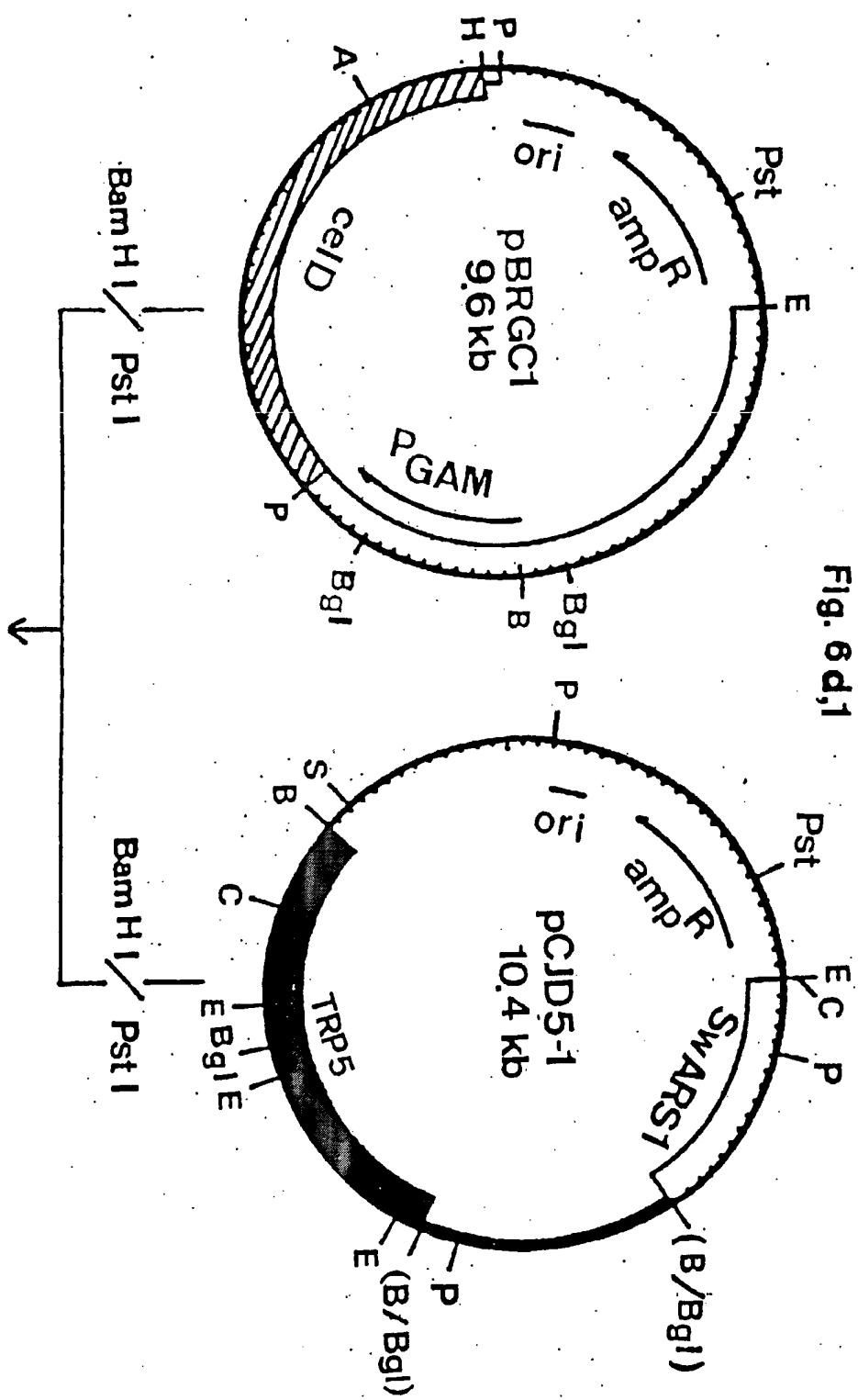


Fig. 6c, 2

Fig. 6d,1



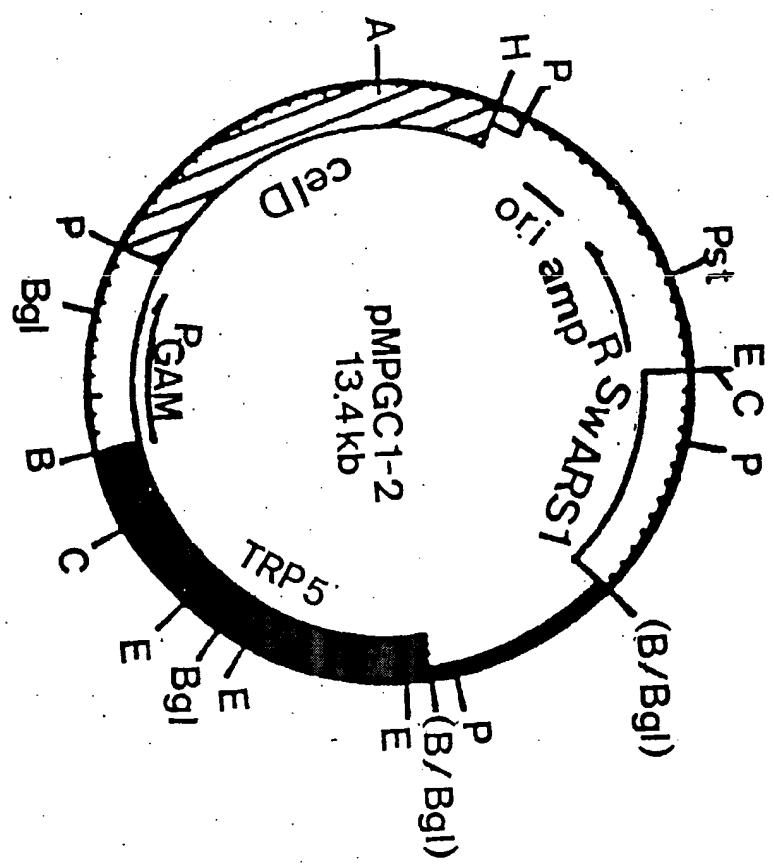


Fig.6d,2

Fig.7a

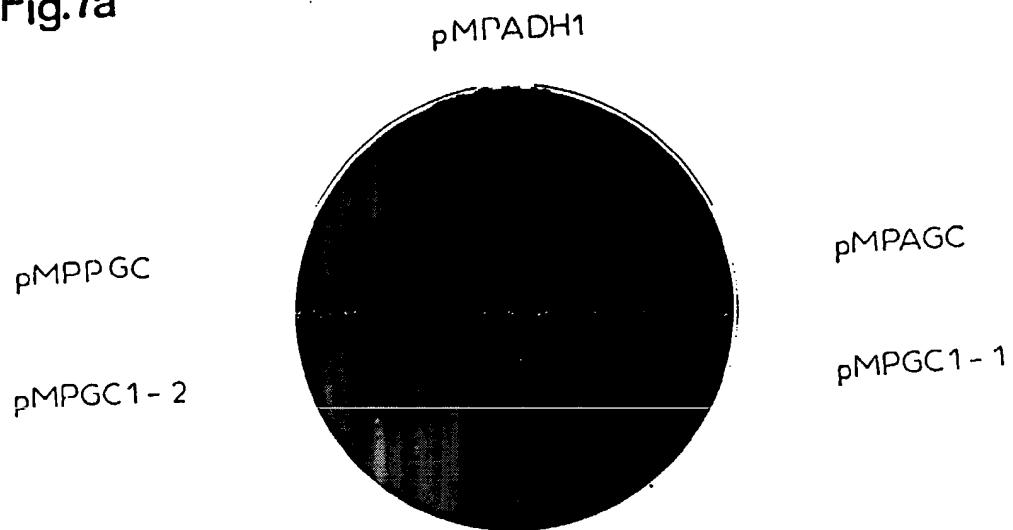
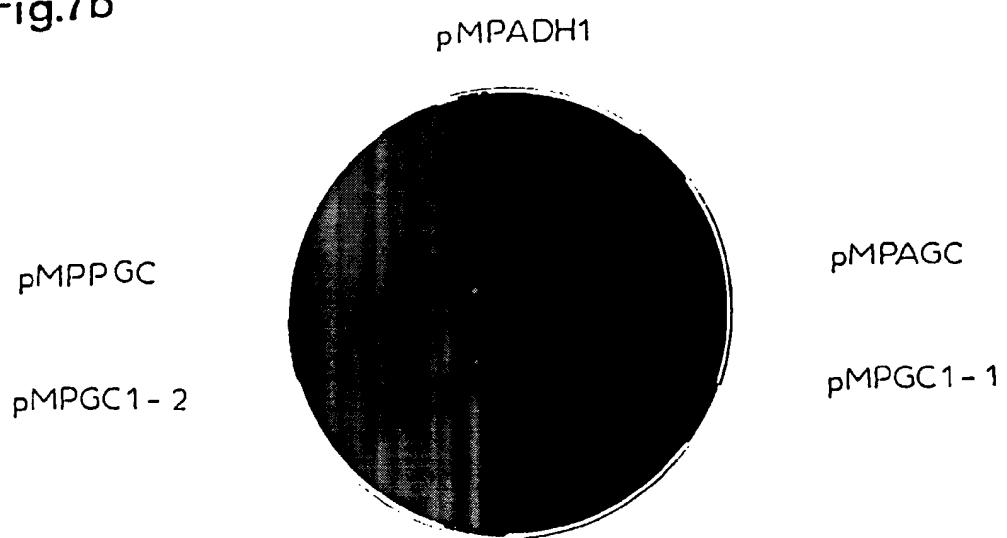
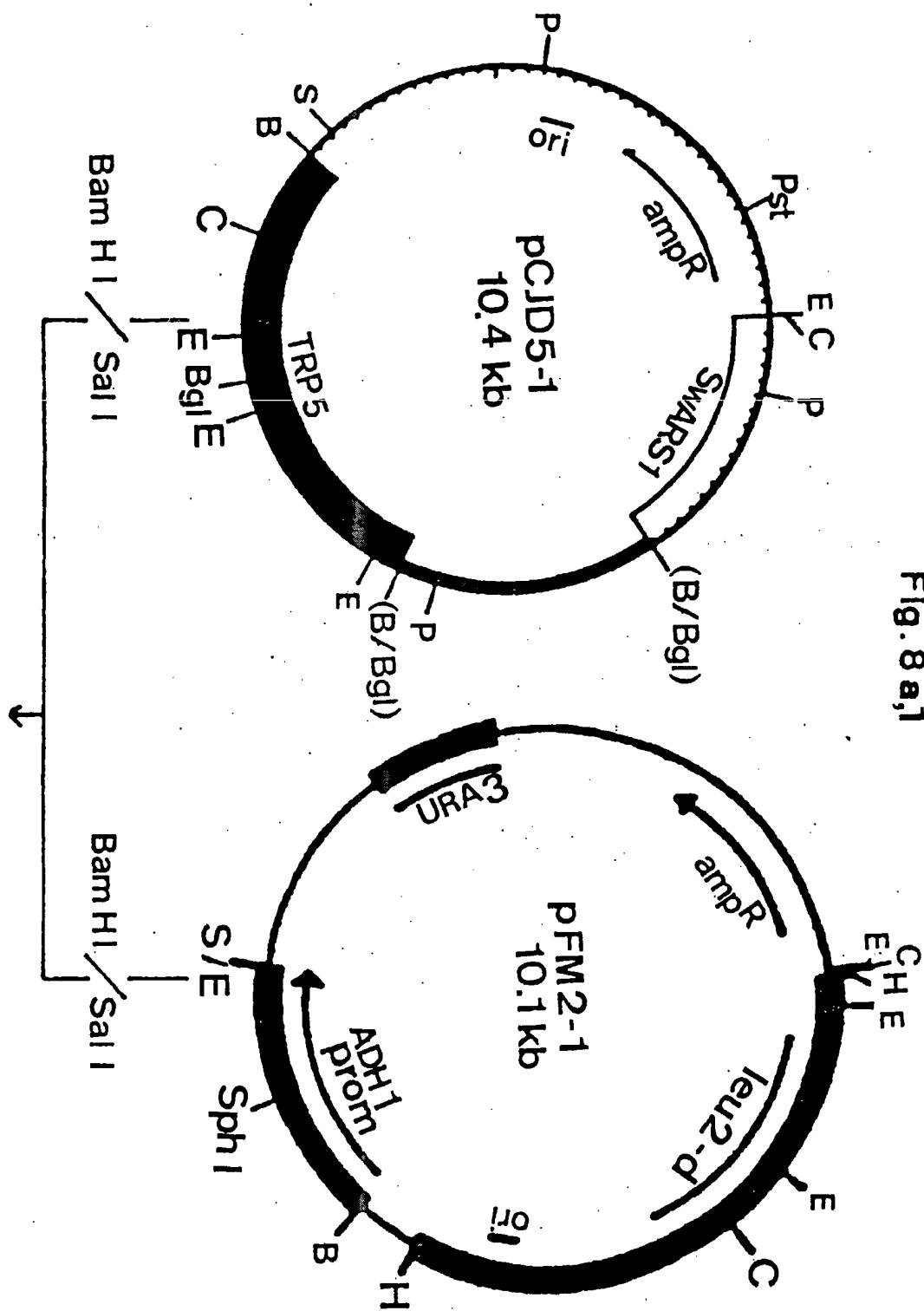


Fig.7b





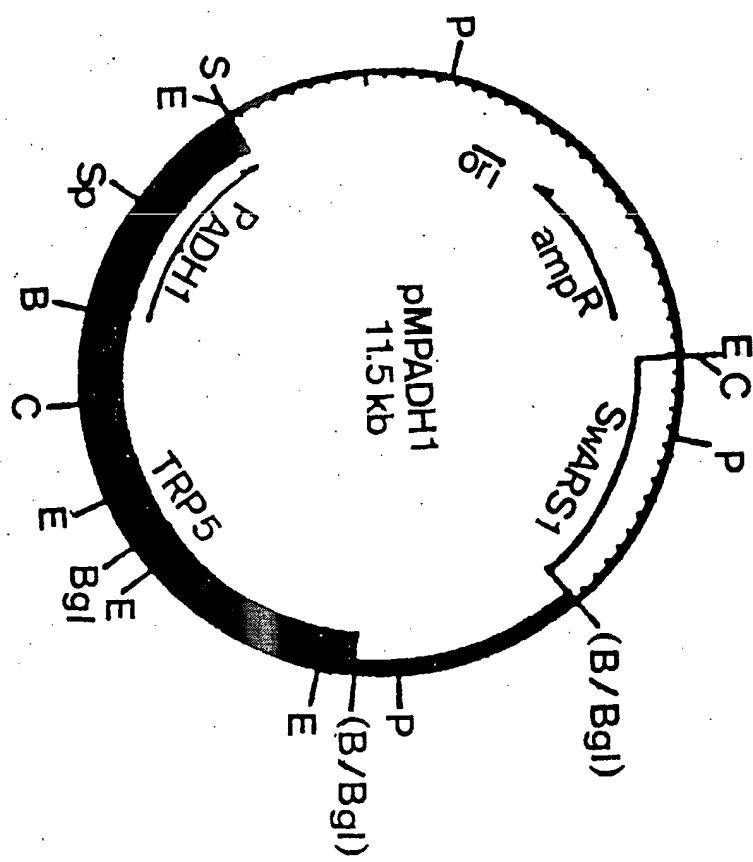
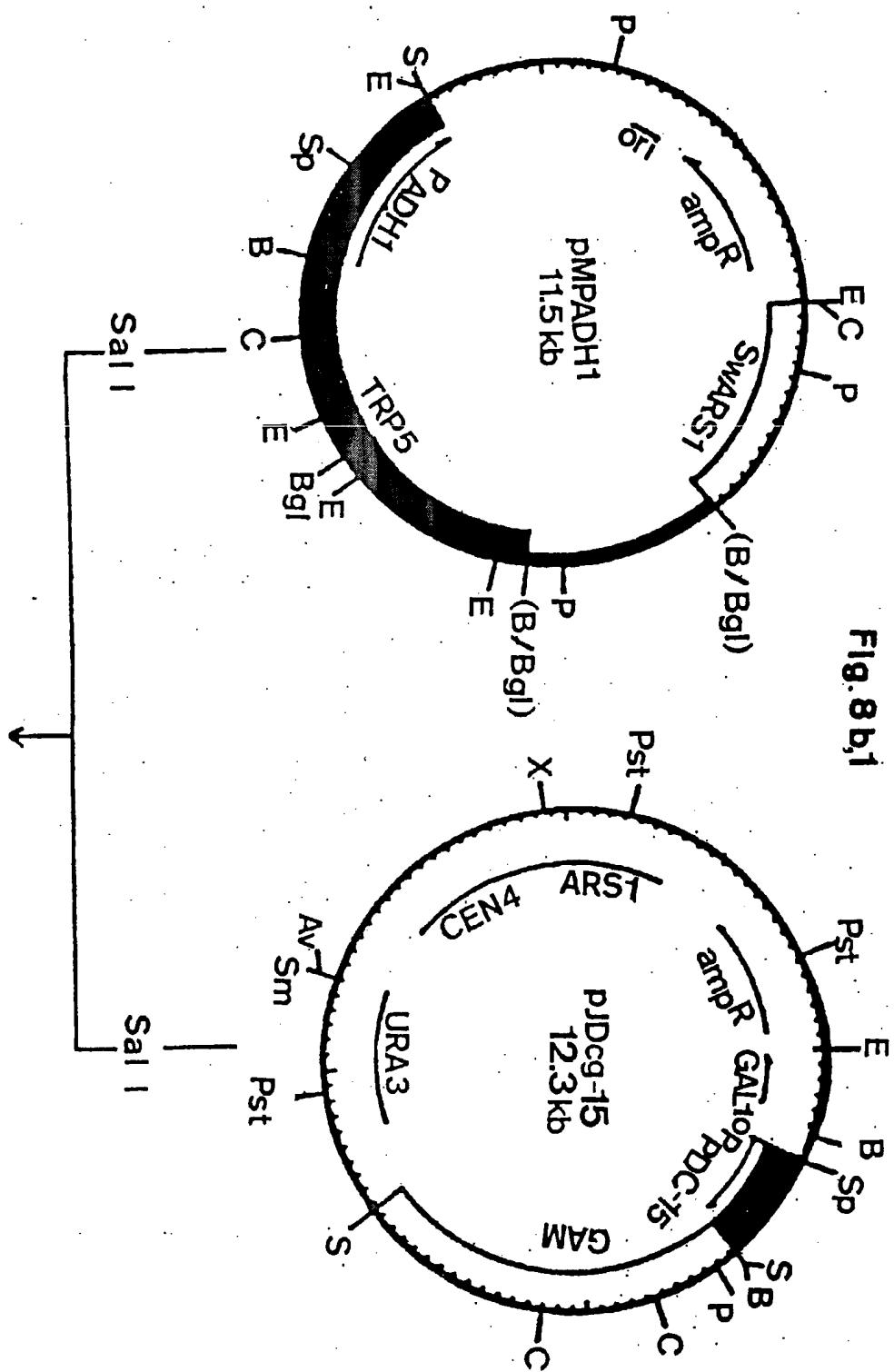


Fig.8a,2



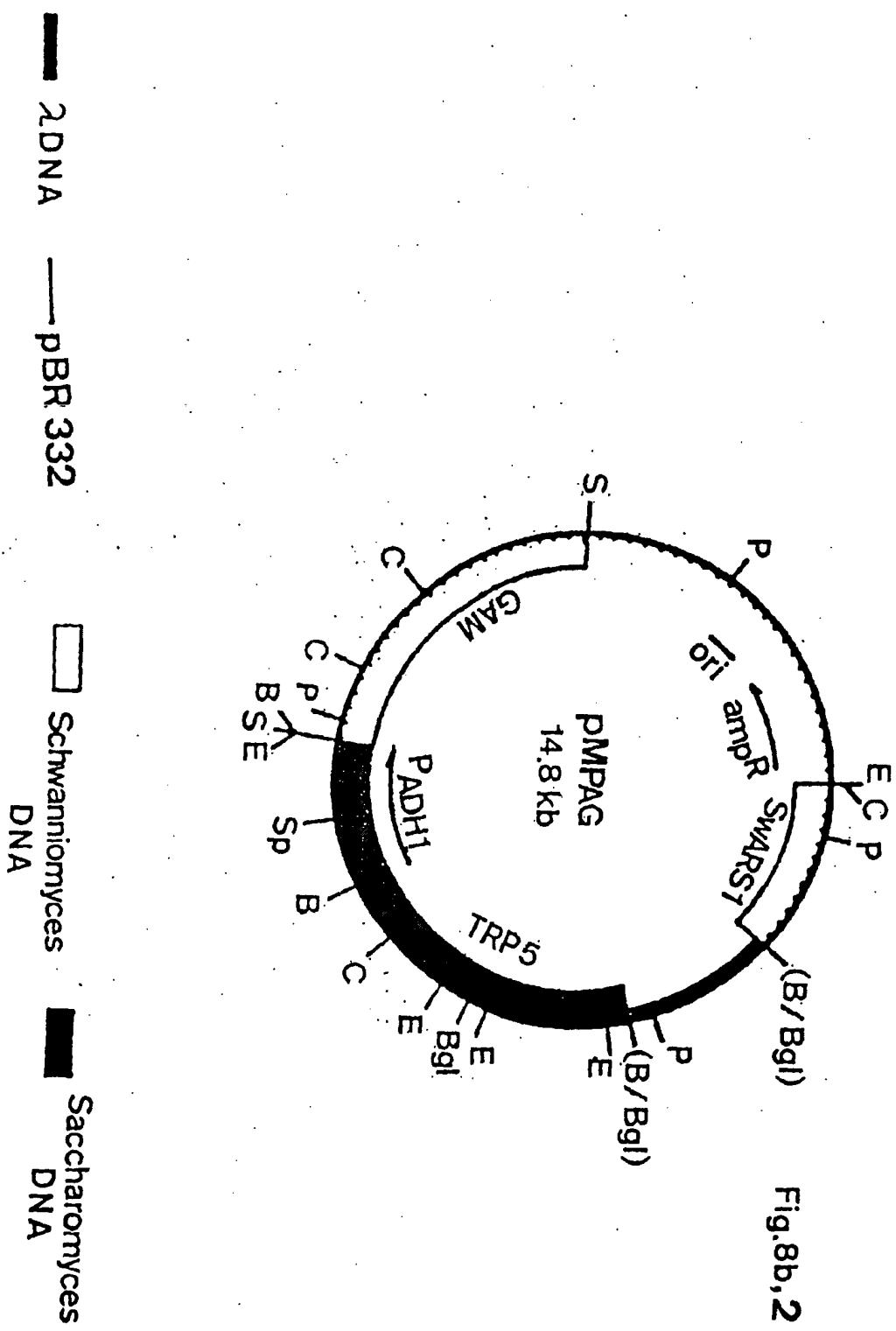


Fig. 8c, t

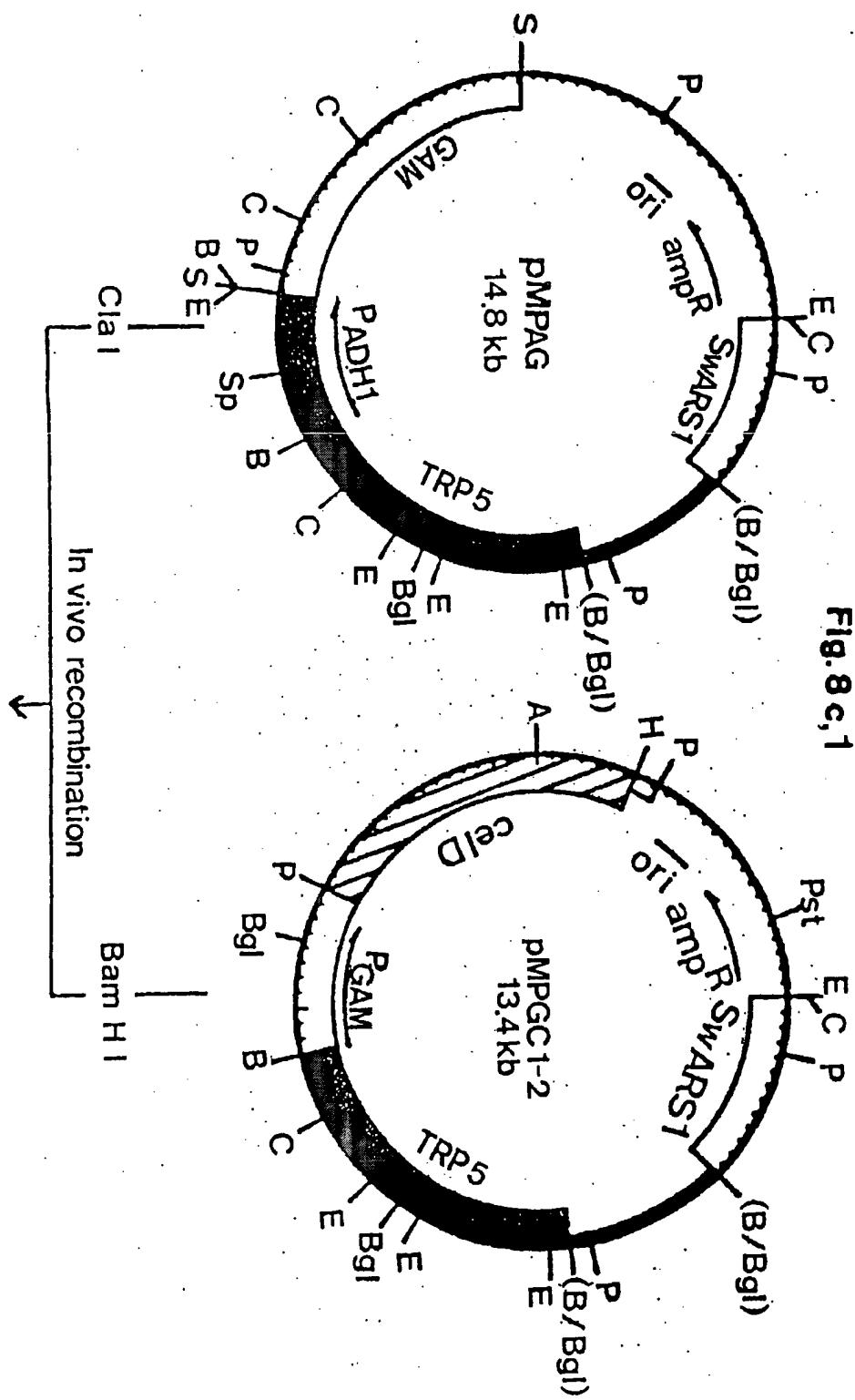
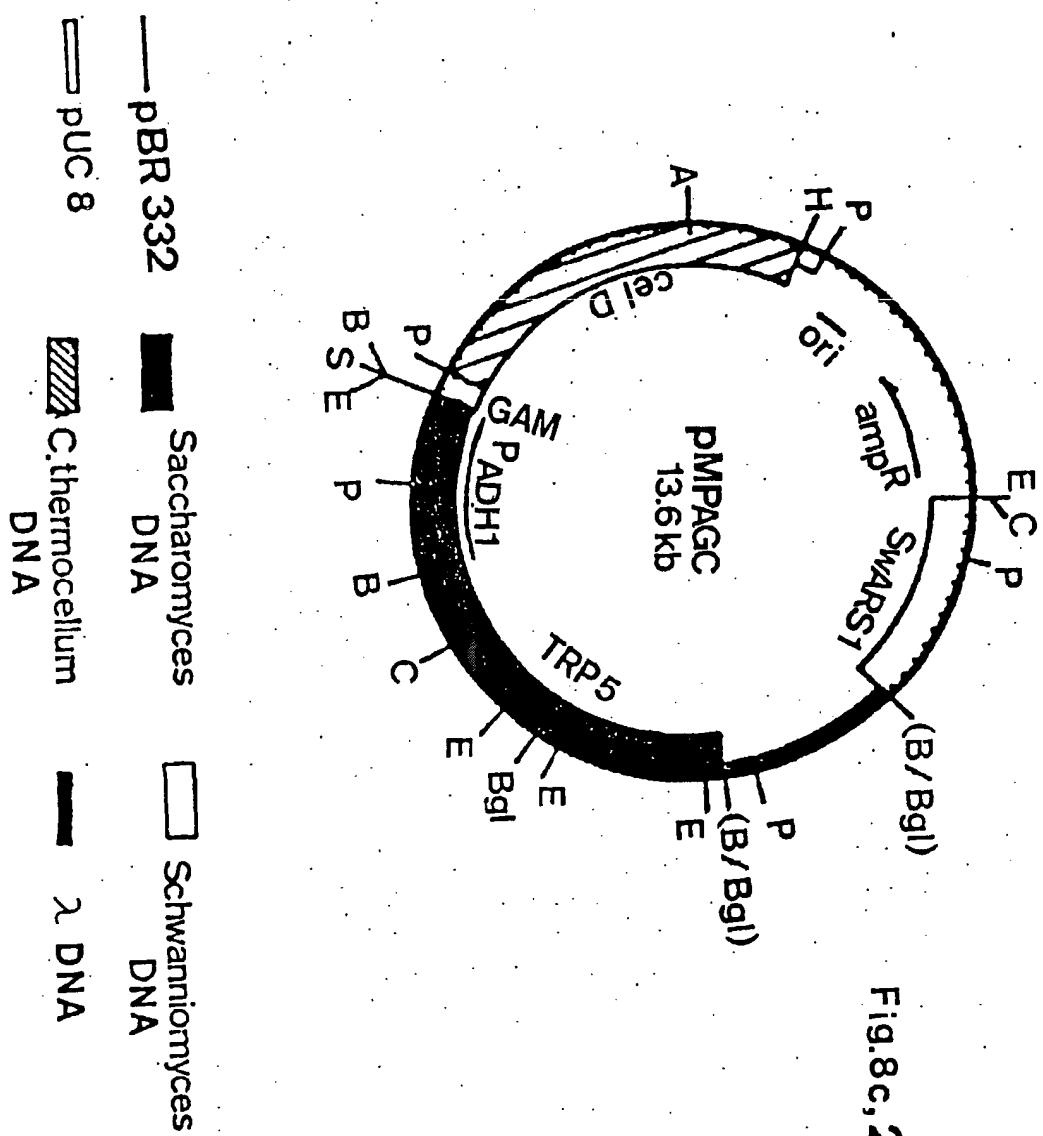


Fig.8c,2



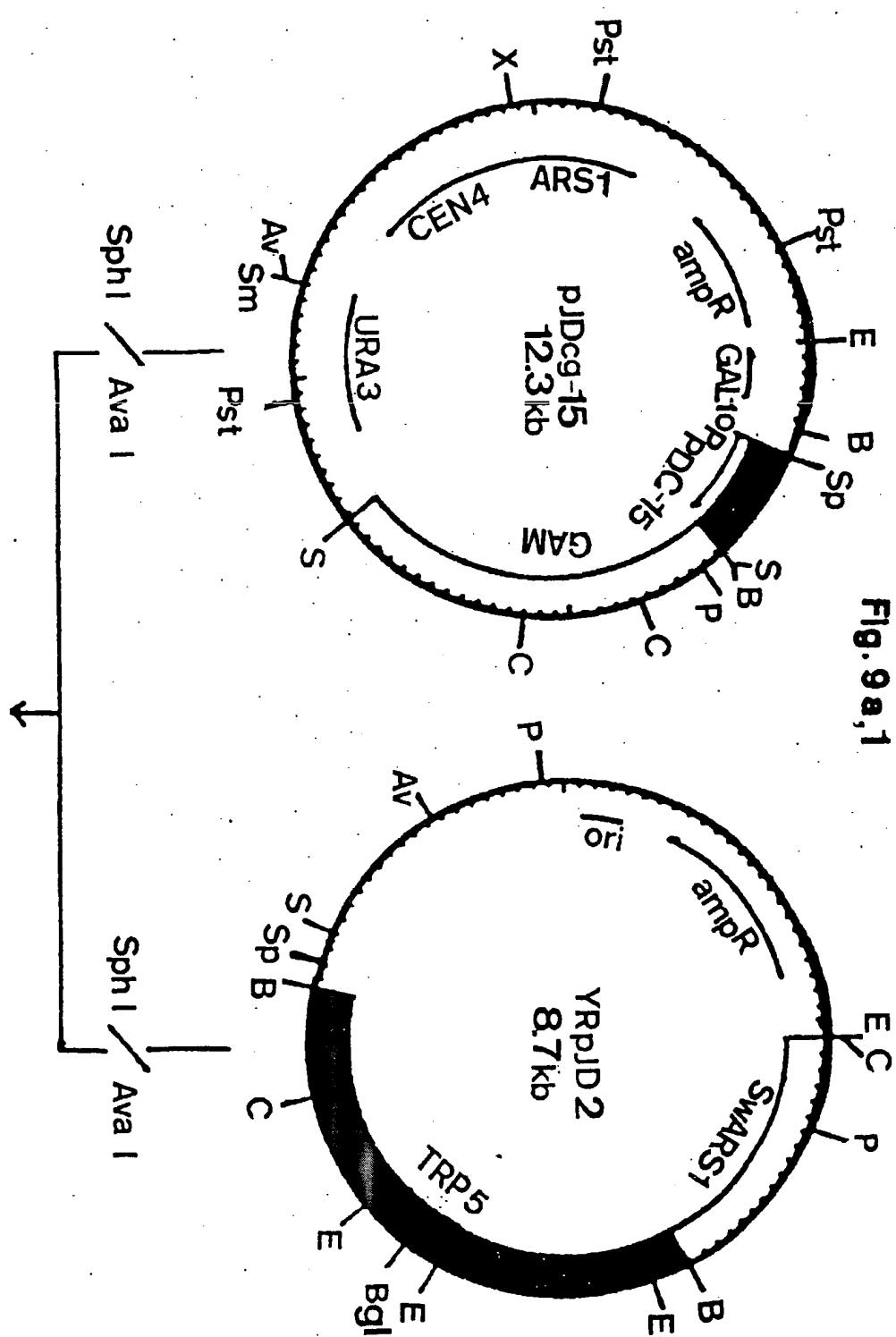


Fig. 9a, 1

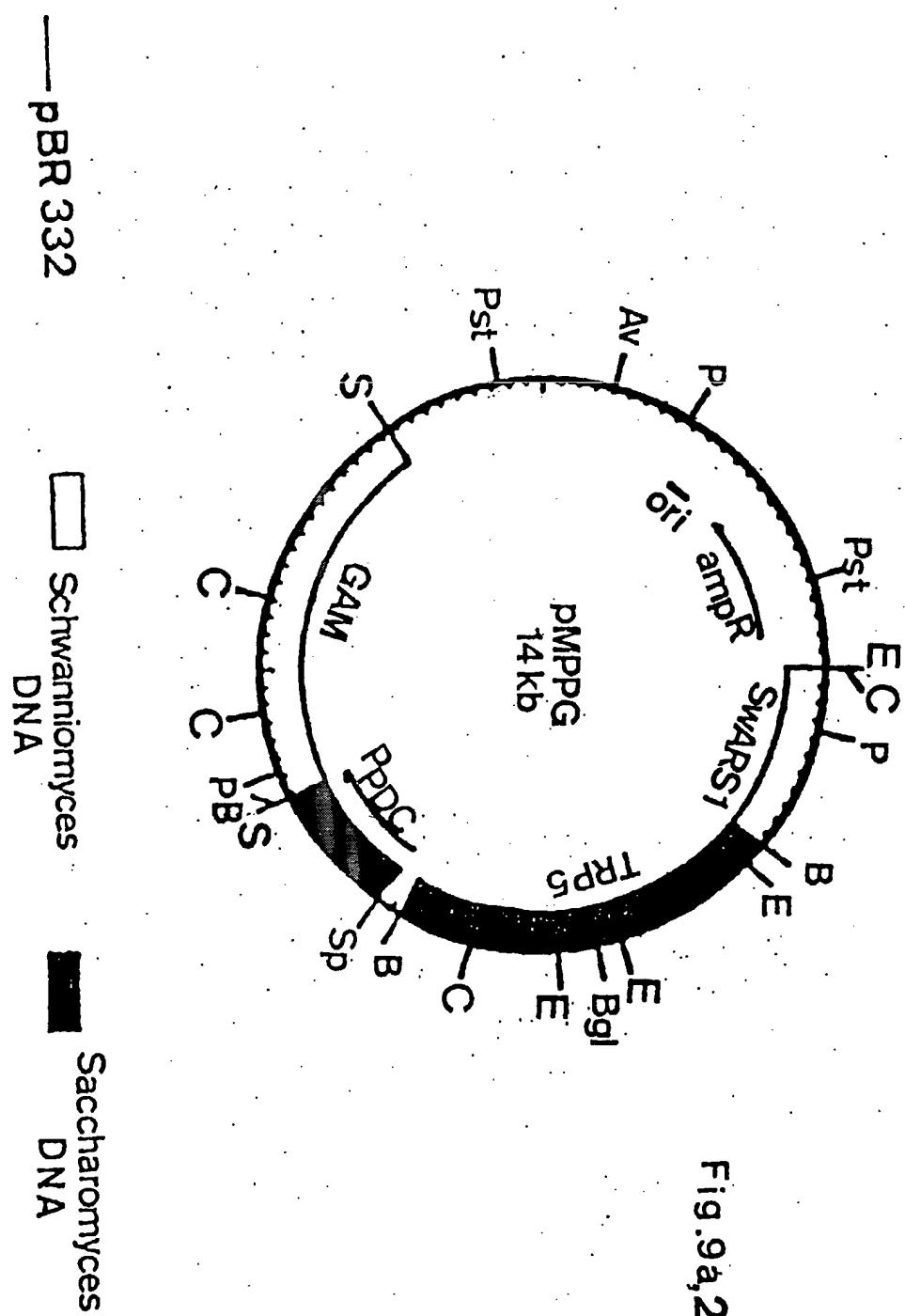
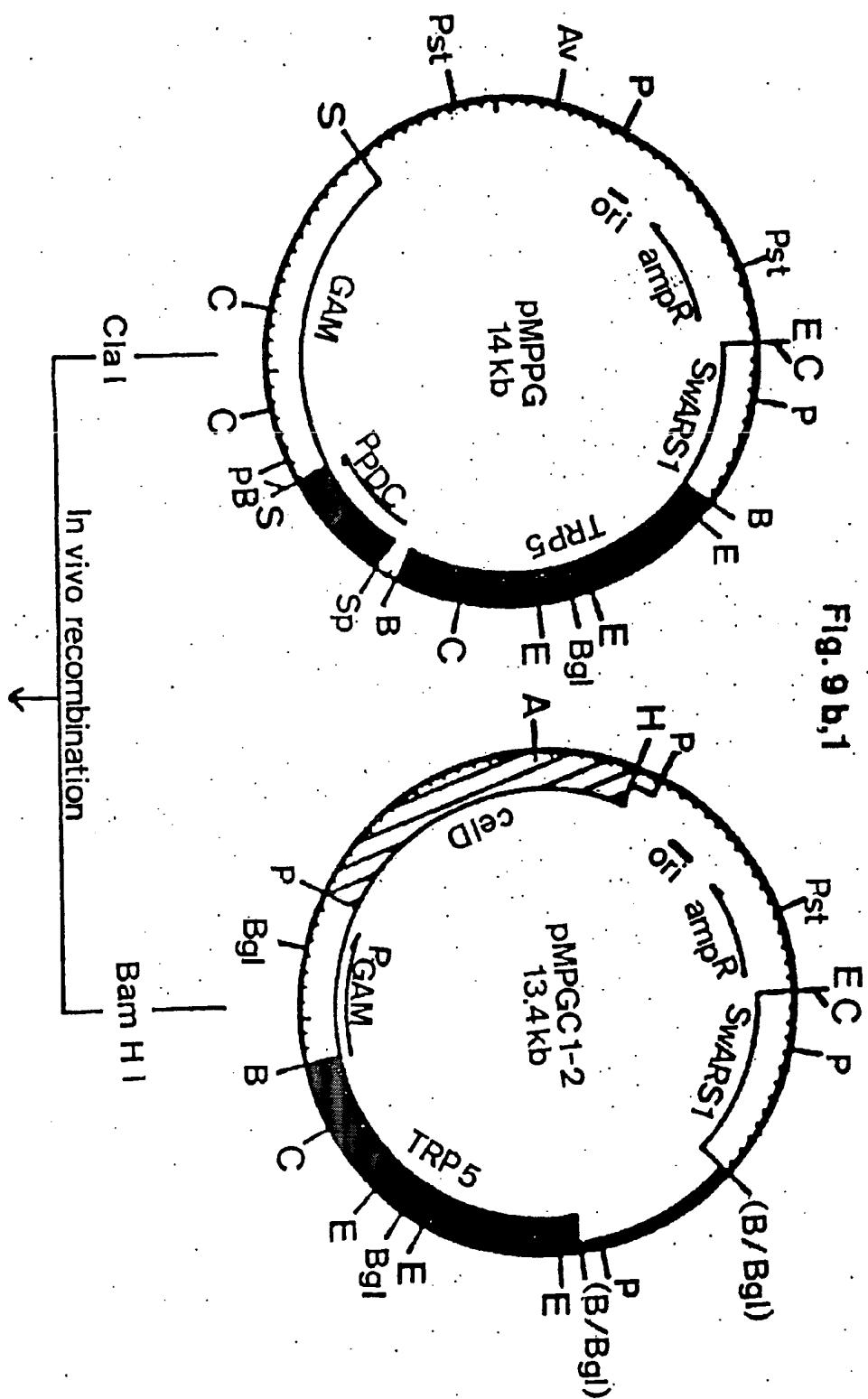


Fig.9a,2

Fig. 9 b,1



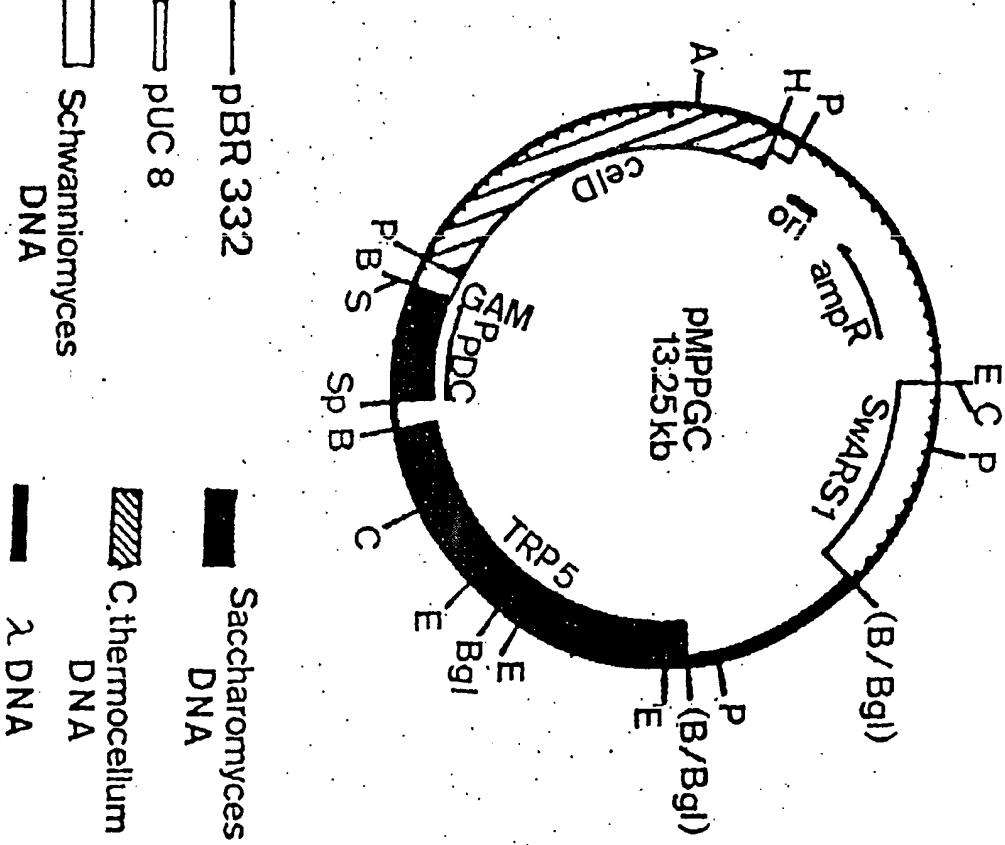


Fig.9b,2

Fig. 10,1

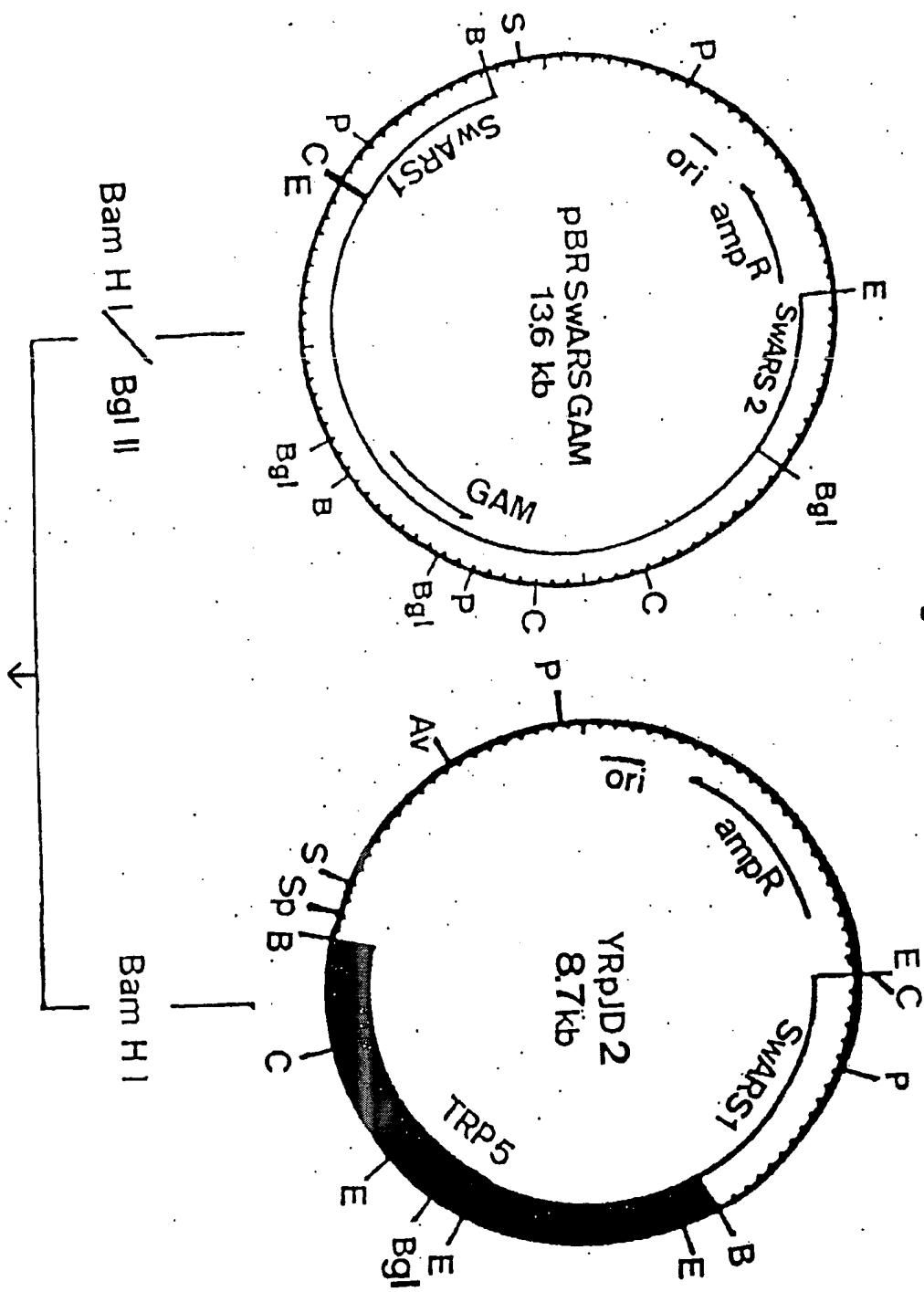


Fig.10,2

